

192663

STIC-Biotech/ChemLib

From: Ramirez, Delia
Sent: Monday, June 12, 2006 3:12 PM
To: STIC-Biotech/ChemLib
Subject: 10/612779

Hi,

I would like to request the following searches: SEQ ID NO:6 and 30 in the protein databases (commercial & interference).

Please provide a hard copy of this search.

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

78201

CRFE

BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE IS BLANK

Thu Jun 15 13:22:24 2006

us-10-612-779-6.rag

Page 1

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:14:44 ; Search time 286.262 Seconds
(without alignments)
972.693 Million cell updates/sec

Title: US-10-612-779-6

Perfect score: 3089

Sequence: 1 MCGIVGATAQRDVAEILLEG.....LIKGTVDVQPRNLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 8:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*
- 10: Geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3089	100.0	609	3	AAY58824 E. coli g
2	3089	100.0	609	8	ADI38861 Mutant gl
3	3089	100.0	609	8	ADU00592 Amino aci
4	3072	99.4	609	3	AAY58827 E. coli g
5	3072	99.4	609	3	AAY58826 E. coli g
6	3072	99.4	609	8	ADI38865 Mutant gl
7	3072	99.4	609	8	ADI38867 Mutant gl
8	3068	99.3	609	3	AAY58822 E. coli g
9	3068	99.3	609	4	AAY34806 E. coli c
10	3068	99.3	609	6	ABU28829 Protein e
11	3068	99.3	609	8	ADI38857 Glucosami
12	3068	99.3	609	8	ADSA45181 Bacterial
13	3068	99.3	609	8	ADU00590 Amino aci
14	3063	99.2	608	9	ADW23840 Novel hum
15	3061	99.1	609	3	AAY58825 E. coli g
16	3061	99.1	609	8	ADI38863 Mutant gl
17	3053	98.8	609	3	AAY58823 E. coli g
18	3053	98.8	609	8	ADI38859 Mutant gl
19	3043	98.5	609	8	ADI38869 Mutant gl
20	3035	98.3	609	4	AAY38472 Salmonell
21	3035	98.3	609	6	ABU48228 Protein e
22	2932	94.9	609	6	ABU28355 Protein e
23	2862	92.7	609	4	AAY36165 Klebsiell

24	2862	92.7	651	7	ABO64174	ABO64174 Klebsiell
25	2857	92.5	609	6	ABU31411	ABU31411 Protein e
26	2700.5	87.4	608	6	ABU45041	ABU45041 Protein e
27	2668	86.4	609	6	ABU50609	ABU50609 Protein e
28	2647	85.7	609	8	ADS42812	ADS42812 Bacterial
29	2605	84.3	609	8	ADN17785	ADN17785 Bacterial
30	2558	82.8	610	6	ABM70565	ABM70565 Phototrab
31	2498.5	80.9	622	7	ADF07237	ADF07237 Bacterial
32	2493.5	80.7	608	6	ABU41104	ABU41104 Protein e
33	2256.5	73.0	610	4	AAU35425	AAU35425 Haemophil
34	2256.5	73.0	610	6	ABU30239	ABU30239 Protein e
35	2256.5	73.0	610	6	AAE30459	AAE30459 Haemophil
36	2256.5	73.0	610	10	AAE31344	AAE31344 Haemophil
37	2250.5	72.9	610	6	ABU49139	ABU49139 Protein e
38	2188.5	70.8	610	6	ABU39464	ABU39464 Protein e
39	1993	64.5	611	6	ABU39730	ABU39730 Protein e
40	1990.5	64.4	610	8	ADS24814	ADS24814 Bacterial
41	1983	64.2	611	4	AAU36518	AAU36518 Pseudomon
42	1983	64.2	611	6	ABU38924	ABU38924 Protein e
43	1983	64.2	616	7	ABO70418	ABO70418 Pseudomon
44	1975	63.9	611	6	ABU41714	ABU41714 Protein e
45	1797.5	58.2	612	6	ABU17019	ABU17019 Protein e

ALIGNMENTS

RESULT 1
AAY58824
ID AAY58824 standard; protein; 609 AA.
XX
AC AAY58824;
XX
DT 08-MAY-2000 (first entry)
XX
DE E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-54.
XX
KW Glucosamine-6-phosphate synthase; glmS gene; mutant; GlcN6P-S-54;
KW Glucosamine; metabolic engineering; plasmid pKLN23-54; mutein.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Misc-difference 39 /note= "replaces wild-type Ala"
FT Misc-difference 250 /note= "replaces wild-type Arg"
FT Misc-difference 472 /note= "replaces wild-type Gly"
XX
WO200004182-A1.
PN
XX
27-JAN-2000.
PD
XX
15-JUL-1999; 99WO-US015976.
PF
XX
15-JUL-1998; 98US-00115475.
PR
XX
(DCVB-) DCV INC DBA BIO-TECH RESOURCES.
PA
XX
Berry A. Burlingame RP, Millis JR;
XX
WPI; 2000-182441/16.
DR
XX
N-PSDB; AAZ58251.
DR
XX
Fermentation of E. coli having an altered amino acid sugar metabolic
PT pathway to produce glucosamine, especially using novel recombinant
PT variant glucosamine-6-phosphate synthases.
XX
Claim 28; Page 124-126; 150pp; English.
XX
The present sequence is that of a mutant, denoted GlcN6P-S-54, of the
CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.

CC When compared with the wild-type sequence (see AAY58822), the mutant
CC includes A39T, R250C and G472S amino acid substitutions. These
CC alterations are predicted from the mutated glms gene in plasmid pKLN23-54
CC (see AAZ58251). The invention provides methods for the overproduction of
CC glucosamine by fermentation using a genetically engineered microorganism,
CC especially E. coli, that includes a modified GlcN6P synthase. Production
CC of the glucosamine by recombinant strain 2123-54 (pKLN23-54) was
CC significantly increased when compared to a strain expressing wild-type
CC GlcN6P synthase owing to reduced product inhibition. Glucosamine
CC concentrations of over 12 g/l have been obtained

XX SQ Sequence 609 AA;

Query Match 100.0%; Score 3089; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 2,7e-271;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKVMQAQAAE 60
Qy 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYCHYMOKEIYEQPNAINKTTLTGRI SHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIYEQPNAINKTTLTGRI SHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLA 360
Qy 361 RLSKELGVLGSLAICNVPGSSIVRESDLAMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGVLGSLAICNVPGSSIVRESDLAMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSQDKRIEALAEFSDKHAFLSRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSQDKRIEALAEFSDKHAFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYIHAEVAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNTEEVRA 540
Db 481 LEGALKLKEISYIHAEVAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNTEEVRA 540
Qy 541 GGQLVFDADQAGFVSSDNMHIEMPHVEEVIAPFTYVLPQLLAHYHVALIKGTDVQDP 600
Db 541 GGQLVFDADQAGFVSSDNMHIEMPHVEEVIAPFTYVLPQLLAHYHVALIKGTDVQDP 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

ADI38861
ADI38861 standard; protein; 609 AA.
XX
AC ADI38861;
XX
DT 15-APR-2004 (first entry)
XX
DE Mutant glucosamine-6-phosphate synthase, glms*54, SEQ ID 6.
XX
, KW Glucosamine; N-acetylglucosamine; fermentation;

KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.
OS Escherichia coli.
OS Synthetic.
XX
XX WO2004003175-A2.
XX
XX 08-JAN-2004.
XX
XX 01-JUL-2003; 2003WO-US020925.
XX
XX 01-JUL-2002; 2002US-0393348P.
XX
XX (ARKI-) ARKION LIFE SCI LLC.
XX
XX Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
XX Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
XX WPI; 2004-203380/19.
XX
XX N-PSDB; ADI38860.
XX
XX Producing glucosamine or N-acetylglucosamine by fermentation involves
XX culturing microorganism comprising glucosamine-6-phosphate
XX acetyltransferase, in fermentation medium, and collecting product.
XX
XX Claim 15; SEQ ID NO 6; 327pp; English.
XX
XX The present invention relates to a method (M1) for producing glucosamine
XX and N-acetylglucosamine by fermentation. The method comprises (a)
XX culturing in a fermentation medium a microorganism (I) which comprises
XX endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
XX modification that increases the activity of (II), glucosamine-6-phosphate
XX synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
XX the activity of (IV) and increases the activity of glucosamine-1-
XX phosphate N-acetyltransferase (V), and (b) and collecting the product,
XX which is chosen from the group consisting of glucosamine-6-phosphate,
XX glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
XX acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
XX sequence was used to illustrate the method of the invention.

SQ Sequence 609 AA;

Query Match 100.0%; Score 3089; DB 8; Length 609;
Best Local Similarity 100.0%; Pred. No. 2,7e-271;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKVMQAQAAE 60
Qy 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYCHYMOKEIYEQPNAINKTTLTGRI SHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIYEQPNAINKTTLTGRI SHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLA 360

QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
 Db |||||
 QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
 Db |||||
 QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPTA 480
 Db |||||
 QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLEKKNIEBVRAR 540
 Db |||||
 QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDPR 600
 Db |||||
 QY 601 NLAKSVTVE 609
 Db |||||
 QY 601 NLAKSVTVE 609
 Db |||||
 RESULT 3
 ID ADU00592
 XX ADU00592 standard; protein; 609 AA.
 AC ADU00592;
 XX
 DT 13-JAN-2005 (first entry)
 XX
 DE Amino acid sequence of a mutant glucosamine-6-phosphate synthase.
 KW chitin; chitosan; fermentation;
 KW glucamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;
 KW glucosamine-6-phosphate synthase; GFA1 gene; glms gene;
 KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;
 KW chitin deacetylase; CDA2 gene;
 KW N-acetylglucosamine-6-phosphate deacetylase;
 KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;
 KW enzyme; glms*54.
 XX
 OS Escherichia coli.
 XX
 PN WO2004092391-A2.
 XX
 PD 28-OCT-2004.
 XX
 PF 12-APR-2004; 2004WO-US011286.
 XX
 PR 11-APR-2003; 2003US-0462087P.
 XX
 PA (ARKI-) ARKION LIFE SCI LLC.
 XX
 PI Deng M, McMullin TW, Grund AD;
 XX
 DR WPI: 2004-766880/75.
 DR N-PSDB; ADU00591.
 XX
 PT Producing chitin or chitosan, involves culturing microorganism with
 PT genetic modification that results in increase in activity of glutamine-
 PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce
 PT chitin or chitosan.
 XX
 PS Example 1; SEQ ID NO 25; 161pp; English.
 CC
 CC The specification describes a method for producing chitin or chitosan by
 CC a fermentation process. The method involves culturing in a fermentation
 CC medium a microorganism which comprises one or more genetic modifications
 CC that result in an increase in the activity of glutamine-fructose-6-
 CC phosphate amidotransferase (also known as glucosamine synthetase and
 CC glucosamine-6-phosphate synthase, and encoded by the GFA1 eukaryotic gene
 CC and the glms bacterial gene), glucosamine-6-phosphate acetyltransferase
 CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1

CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-
 CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and
 CC chitosanase, and collecting the chitin or chitosan. The method is useful
 CC for producing chitin or chitosan by utilizing microorganisms such as
 CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and
 CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.
 CC cerevisiae, A. niger or A. nidulans. The method enables high quantities
 CC of chitin and chitosan to be produced cost effectively. The present
 CC sequence represents a mutant glucosamine-6-phosphate synthase, encoded by
 CC glms*54. This enzyme is resistant to feedback inhibition by glucosamine-6-
 CC -phosphate. The polynucleotide was used to transform yeast for use in the
 CC method of the invention.
 XX
 SQ Sequence 609 AA;
 Query Match 100.0%; Score 3089; DB 8; Length 609;
 Best Local Similarity 100.0%; Pred. No. 2.7e-271;
 Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDVSAGLAVVDTGEGHMTLRRLRGKQVQLAAAE 60
 Db |||||
 QY 61 EHLPGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIENHEPLREELKARGYTFV 120
 Db |||||
 QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
 Db |||||
 QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180
 QY 181 IGLNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGARVKQDIESNQ 240
 Db |||||
 QY 181 IGLNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGARVKQDIESNQ 240
 QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
 Db |||||
 QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
 QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAGL 360
 Db |||||
 QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAGL 360
 QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
 Db |||||
 QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
 QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPTA 480
 Db |||||
 QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLEKKNIEBVRAR 540
 Db |||||
 QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDPR 600
 Db |||||
 QY 601 NLAKSVTVE 609
 Db |||||
 QY 601 NLAKSVTVE 609
 Db |||||
 RESULT 4
 ID AAY58827
 XX AAY58827 standard; protein; 609 AA.
 AC AAY58827;
 XX
 DT 08-MAY-2000 (first entry)
 XX
 DE E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-151.

XX KW Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-151;
 KW glucosamine; metabolic engineering; plasmid pKLN23-151; mutein.
 XX OS Escherichia coli.
 XX FH Key Location/Qualifiers
 FT Misc-difference 472 /note= "replaces wild-type Gly"
 FT W0200004182-A1.
 XX PD 27-JAN-2000.
 XX PF 15-JUL-1999; 99WO-US015976.
 XX PR 15-JUL-1998; 98US-00115475.
 XX PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
 XX PI Berry A, Burlingame RP, Millis JR;
 XX WPI; 2000-182441/16.
 XX DR N-PSDB; AAZ58254.
 XX PT Fermentation of E. coli having an altered amino acid sugar metabolic
 PT pathway to produce glucosamine, especially using novel recombinant
 PT variant glucosamine-6-phosphate synthases.
 XX PS Claim 28; Page 143-145; 150pp; English.
 XX CC The present sequence is that of a mutant, denoted GlcN6P-S-151, of the
 CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
 CC When compared with the wild-type sequence (see AAY58822), the mutant
 CC includes a Ser for Gly-472 amino acid substitution. This alteration was
 CC predicted from the mutated glms gene in plasmid pKLN23-151 (see AAZ58254).
 CC The invention provides methods for the overproduction of glucosamine by
 CC fermentation using a genetically engineered microorganism, especially E.
 CC coli, that includes a modified GlcN6P synthase. Production of the
 CC glucosamine by recombinant strain 2123-149 (pKLN23-149) was marginally
 CC increased when compared to a strain expressing wild-type GlcN6P synthase
 XX Sequence 609 AA;
 SQ Query Match 99.4%; Score 3072; DB 3; Length 609;
 Beat Local Similarity 99.7%; Pred. No. 9.7e-270;
 Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGIVGAIQORDVAEILLEGRLREYRGYDSAGLAVDTEGHMTRELKGVQMLAQAAE 60
 DB 1 MCGIVGAIQORDVAEILLEGRLREYRGYDSAGLAVDTEGHMTRELKGVQMLAQAAE 60
 QY 61 EHLPLGGTGIAHTRWATHGESEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
 DB 61 EHLPLGGTGIAHTRWATHGESEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
 QY 121 SETDTEVIAHLVNWELKGGTLEAVLRAIPQLRGAYGTVMDSRPHDTLLAARSGSPV 180
 DB 121 SETDTEVIAHLVNWELKGGTLEAVLRAIPQLRGAYGTVMDSRPHDTLLAARSGSPV 180
 QY 181 IGLGNGENFIASDQALLPVTRRFTFLEEGDIAETTRSVNIFDKTGAEVKRODIESNLQ 240
 DB 181 IGLGNGENFIASDQALLPVTRRFTFLEEGDIAETTRSVNIFDKTGAEVKRODIESNLQ 240
 QY 241 YDAGDKGIYHMQKEIYEQPNAINKTITGRISHGQVDLSLSELPNADELLSKVEHIQILA 300
 DB 241 YDAGDKGIYHMQKEIYEQPNAINKTITGRISHGQVDLSLSELPNADELLSKVEHIQILA 300
 QY 301 CQTSNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMTITLQSGETADTLA 360
 DB 301 CQTSNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMTITLQSGETADTLA 360
 QY 361 RLSKELGYLGSIAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQTLTVLLMLVAKL 420

DB 361 RLSKELGYLGSIAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQTLTVLLMLVAKL 420
 QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLSRGDPYPIA 480
 DB 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLSRGDPYPIA 480
 QY 481 LEGALKKEISYTHAEAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
 DB 481 LEGALKKEISYTHAEAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
 QY 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFTVTPQLLAYHVALIKGTDVDDQPR 600
 DB 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFTVTPQLLAYHVALIKGTDVDDQPR 600
 QY 601 NLAKSVTVE 609
 DB 601 NLAKSVTVE 609
 RESULT 5
 AAY58826
 ID AAY58826 standard; protein; 609 AA.
 AC AAY58826;
 XX 08-MAY-2000 (first entry)
 XX E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-149.
 XX Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-149;
 XX glucosamine; metabolic engineering; plasmid pKLN23-149; mutein.
 XX Escherichia coli.
 XX Key Location/Qualifiers
 FT Misc-difference 472 /note= "replaces wild-type Gly"
 FT W0200004182-A1.
 XX PD 27-JAN-2000.
 XX PF 15-JUL-1999; 99WO-US015976.
 XX PR 15-JUL-1998; 98US-00115475.
 XX PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
 XX Berry A, Burlingame RP, Millis JR;
 XX WPI; 2000-182441/16.
 XX DR N-PSDB; AAZ58253.
 XX PT Fermentation of E. coli having an altered amino acid sugar metabolic
 PT pathway to produce glucosamine, especially using novel recombinant
 PT variant glucosamine-6-phosphate synthases.
 XX PS Claim 28; Page 137-139; 150pp; English.
 XX CC The present sequence is that of a mutant, denoted GlcN6P-S-149, of the
 CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
 CC When compared with the wild-type sequence (see AAY58822), the mutant
 CC includes a Ser for Gly-472 amino acid substitution. This alteration was
 CC predicted from the mutated glms gene in plasmid pKLN23-149 (see AAZ58253).
 CC The invention provides methods for the overproduction of glucosamine by
 CC fermentation using a genetically engineered microorganism, especially E.
 CC coli, that includes a modified GlcN6P synthase. Production of the
 CC glucosamine by recombinant strain 2123-149 (pKLN23-149) was marginally
 CC increased when compared to a strain expressing wild-type GlcN6P synthase
 XX Sequence 609 AA;
 SQ

Query Match	99.4%;	Score 3072;	DB 3;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
QY	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
Db	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
QY	601	NL	AK	SVT	VE 609	
Db	601	NL	AK	SVT	VE 609	
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
QY	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
Db	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
QY	601	NL	AK	SVT	VE 609	
Db	601	NL	AK	SVT	VE 609	
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
QY	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
Db	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
QY	601	NL	AK	SVT	VE 609	
Db	601	NL	AK	SVT	VE 609	
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
QY	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
Db	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
QY	601	NL	AK	SVT	VE 609	
Db	601	NL	AK	SVT	VE 609	
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
QY	241	YD	AGDKGI	YCHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
Db	241	YD	AGDKGI	YRHYMQKEIYEQPNAIKNTLTGRISHGOVDLS	ELGPNADLLSKVEHIQILA 300	
QY	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
Db	301	CG	TSYNSGMVSR	YWFESLAGICDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLA 360	
QY	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db	361	RL	SK	ELG	YLGSLAICNVPGSSLVRES	DALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
QY	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
Db	421	SR	LK	GLD	ASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLSRGDQYPIA 480	
QY	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
Db	481	LE	GALK	KBISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNLEKLSKSNIEVRAR 540		
QY	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
Db	541	GG	QLY	VFAD	QDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600	
QY	601	NL	AK	SVT	VE 609	
Db	601	NL	AK	SVT	VE 609	
us-10-612-779-6.rag						
Query Match	99.4%;	Score 3072;	DB 8;	Length 609;		
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;				
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60	
Db	1	MC	GIVGAI	AQ	RDVAEILLEGRLREYGYDSAGLAVVDAEGHTRRLRLGKVQMLAQAAE 60	
QY	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
Db	61	EHP	HGGT	GI	AHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120	
QY	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	SET	DEVIA	HLVNVNELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
QY	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI FDKTGAEVKRODIESNLQ 240	
Db	181	IG	MG	ENFI	ASDQALLPVTRRFIFLEEGDIAETITRSVNI	

QY 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
 DB 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
 QY 601 NLAKSVTVE 609
 DB 601 NLAKSVTVE 609
 RESULT 7
 ADI38867
 ID ADI38867 standard; protein; 609 AA.
 AC ADI38867;
 DT 15-APR-2004 (first entry)
 DE Mutant glucosamine-6-phosphate synthase, SEQ ID 12.
 KW Glucosamine; N-acetylglucosamine; fermentation;
 KW glucosamine-6-phosphate acetyltransferase;
 KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
 KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;
 KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
 KW N-acetylglucosamine-6-phosphate; enzyme.
 OS Escherichia coli.
 OS Synthetic.
 PN WO2004/003175-A2.
 XX 08-JAN-2004.
 XX 01-JUL-2003; 2003WO-US020925.
 PR 01-JUL-2002; 2002US-0393348P.
 PA (ARKI-) ARKION LIFE SCI LLC.
 PI Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
 PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
 DR WPI: 2004-203380/19.
 DR N-PSDB; ADI38866.
 XX
 PT Producing glucosamine or N-acetylglucosamine by fermentation involves
 PT culturing microorganism comprising glucosamine-6-phosphate
 PT acetyltransferase, in fermentation medium, and collecting product.
 XX
 PS Claim 15; SEQ ID NO 12; 327pp; English.
 CC The present invention relates to a method (M1) for producing glucosamine
 CC and N-acetylglucosamine by fermentation. The method comprises (a)
 CC culturing in a fermentation medium a microorganism (I) which comprises
 CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
 CC modification that increases the activity of (II), glucosamine-6-phosphate
 CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
 CC the activity of (IV) and increases the activity of glucosamine-1
 CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
 CC which is chosen from the group consisting of glucosamine-6-phosphate, N-
 CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
 CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
 CC sequence was used to illustrate the method of the invention.
 XX
 SQ Sequence 609 AA;
 Query Match 99.4%; Score 3072; DB 8; Length 609;
 Best Local Similarity 99.7%; Pred. No. 9.7e-270;
 Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGIVGAIQAQRDVAIILLEGRLRLEYRGVDSAGLAVDTEGHMTLRRLRGKVMQAQAE 60
 DB 1 MCGIVGAIQAQRDVAIILLEGRLRLEYRGVDSAGLAVDTEGHMTLRRLRGKVMQAQAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPUREELKARGYTFV 120
 DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPUREELKARGYTFV 120
 QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLAAARSGSLV 180
 DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLAAARSGSLV 180
 QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKRDQIESNLQ 240
 DB 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKRDQIESNLQ 240
 QY 241 YDAGDKGIYHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300
 DB 241 YDAGDKGIYHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300
 QY 301 CGTSVNSGMVSRYPFESLAGIPCDVEIASEFRYKSAVRRNSLMTLTSQSGETADTLA 360
 DB 301 CGTSVNSGMVSRYPFESLAGIPCDVEIASEFRYKSAVRRNSLMTLTSQSGETADTLA 360
 QY 361 RLSKELGYLGLSLAICNVPSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420
 DB 361 RLSKELGYLGLSLAICNVPSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420
 QY 421 SRLKGLDASIEHDIVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHHLFLSRGQYPIA 480
 DB 421 SRLKGLDASIEHDIVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHHLFLSRGQYPIA 480
 QY 481 LEGALKUKETSITHAEAYAGELKHGLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
 DB 481 LEGALKUKETSITHAEAYAGELKHGLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
 QY 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
 DB 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
 QY 601 NLAKSVTVE 609
 DB 601 NLAKSVTVE 609
 RESULT 8
 AAY58822
 ID AAY58822 standard; protein; 609 AA.
 AC AAY58822;
 XX 08-MAY-2000 (first entry)
 DE E. coli glucosamine-6-phosphate synthase.
 KW Glucosamine-6-phosphate synthase; GlcN6P; glucosamine;
 KW metabolic engineering; plasmid pKLN23-28.
 OS Escherichia coli.
 PN WO200004182-A1.
 XX 27-JAN-2000.
 PF 15-JUL-1999; 99WO-US015976.
 XX 15-JUL-1998; 98US-00115475.
 PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
 PI Berry A, Burlingame RP, Millis JR;
 DR WPI; 2000-182441/16.
 XX N-PSDB; AAZ58249.
 PT Fermentation of E. coli having an altered amino acid sugar metabolic

PT pathway to produce glucosamine, especially using novel recombinant
PT variant glucosamine-6-phosphate synthases.
XX
PS Claim 15; Page 111-113; 150pp; English.
XX
CC The present sequence is that of wild-type glucosamine-6-phosphate
CC synthase (GlcN6P synthase) of *Escherichia coli* strain W3110. The sequence
CC is predicted from the isolated glms gene (see AA258249) in plasmid pKLN23
CC -28. Recombinant nucleic acids encoding GlcN6P synthase are claimed, and
CC are used for the expression of the enzyme in host microbial strains,
CC especially *E. coli*, for use in the production of glucosamine. The
CC invention also provides methods for the overproduction of glucosamine
CC using a genetically engineered microorganism that encodes a GlcN6P
CC synthase modified to increase its activity (see also AA58823-27).
CC Preferred modifications comprise amino acid deletion, insertion,
CC inversion, derivatisation or substitution, especially 14T, 1272T, S240P,
CC A39T, R250C, G472S and L469P substitution, of the present sequence
XX
SQ Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 3; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-269;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRRGKQVQLAAAE 60
DB 1 MCGIVGAIAQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRRGKQVQLAAAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHVVVHNGIIENHPELREELKARGTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHVVVHNGIIENHPELREELKARGTFV 120

QY 121 SETDTEVIAHLVNWELKQGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
DB 121 SETDTEVIAHLVNWELKQGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180

QY 181 IGLGMENFTASDQALLPVTFRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFTASDQALLPVTFRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240

QY 241 YDAGDKGIYCHYQKEIYEOPNAIKNTLTGRISHGQVDLSSELGPNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYQKEIYEOPNAIKNTLTGRISHGQVDLSSELGPNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLAGL 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLAGL 360

QY 361 RLSKELGYLGLSLAICNVPSSLVRESDLAMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
DB 361 RLSKELGYLGLSLAICNVPSSLVRESDLAMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFLSRGDQYPIA 480
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFLSRGDQYPIA 480

QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLKSNIEVRAR 540

QY 541 GGQLYVPADODAGFVSDNNHHIEMPHVEEVIAPIFYTVPLQLLAVHVALIKGTDVDDPR 600
DB 541 GGQLYVPADODAGFVSDNNHHIEMPHVEEVIAPIFYTVPLQLLAVHVALIKGTDVDDPR 600

QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 9
AAU34806
ID AAU34806 standard; protein; 609 AA.
XX

AC AAU34806;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #387.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS *Escherichia coli*.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52665.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
PT
XX Example 3; SEQ ID NO 10399; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-269;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRRGKQVQLAAAE 60
DB 1 MCGIVGAIAQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRRGKQVQLAAAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHVVVHNGIIENHPELREELKARGTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHVVVHNGIIENHPELREELKARGTFV 120

QY 121 SETDTEVIAHLVNWELKQGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
DB 121 SETDTEVIAHLVNWELKQGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180

Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDEVEIASFEPRYKSAVRNSLMTLSQSGTADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDEVEIASFEPRYKSAVRNSLMTLSQSGTADTLAGL 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVGLOALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDIHVGLOALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQOYPIA 480
Qy 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVWAPNNLEKLSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVWAPNNLEKLSNIEVRAR 540
Qy 541 GGQLYVADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
Db 541 GGQLYVADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 10

ABU28829
ID ABU28829 standard; protein; 609 AA.
XX AC ABU28829;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #14356.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Escherichia coli.
XX FN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX PI WPI: 2003-029926/02.
XX DR N-PSDB; AC32699.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 56753; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 6; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-269;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDTEGHMTLRRLRGVQLAAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDTEGHMTLRRLRGVQLAAAE 60
Qy 61 EPHLHGTTGIAHTWATHGEPSEVNAHPHVEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EPHLHGTTGIAHTWATHGEPSEVNAHPHVEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDEVEIASFEPRYKSAVRNSLMTLSQSGTADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDEVEIASFEPRYKSAVRNSLMTLSQSGTADTLAGL 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVGLOALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDIHVGLOALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQOYPIA 480
Qy 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVWAPNNLEKLSNIEVRAR 540

Db 481 LEGALKKEISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKKSIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
ADI38857
ID ADI38857 standard; protein; 609 AA.
AC ADI38857;
XX
XX
DT 15-APR-2004 (first entry)
DE Glucosamine-6-phosphate synthase, glmS, SEQ ID 2.
XX
KW Glucosamine; N-acetylglucosamine; fermentation;
KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW Glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW Glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.
XX Escherichia coli.
XX
XX WO2004003175-A2.
XX
XX 08-JAN-2004.
PF 01-JUL-2003; 2003WO-US020925.
XX
PR 01-JUL-2002; 2002US-0393348P.
XX
XX (ARKI-) ARKION LIFE SCI LLC.
PI Deng M, Angerer JD, Cyton D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
DR WPI; 2004-203380/19.
DR N-PSDB; ADI38856.
XX
XX Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
XX
XX Claim 15; SEQ ID NO 2; 327pp; English.
XX
XX The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate
CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
CC the activity of (IV) and increases the activity of glucosamine-1
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
CC which is chosen from the group consisting of glucosamine-6-phosphate,
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
CC sequence was used to illustrate the method of the invention.
XX
SQ Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 8; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-269;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIQRDVAEILLEGRLRYGYSAGLAVVDTEGHMTRLRRLKGVQMLQAQAE 60

Db 1 MCGIVGAIQRDVAEILLEGRLRYGYSAGLAVVDTEGHMTRLRRLKGVQMLQAQAE 60
QY 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLAARSQPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLAARSQPLV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEBGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEBGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360
QY 361 RLKSELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTOLTLLMLVAKL 420
Db 361 RLKSELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTOLTLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFSLRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFSLRGDQYPIA 480
QY 481 LEGALKKEISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKKSIEEVRAR 540
Db 481 LEGALKKEISYTHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKKSIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 12
ADI38857
ID ADS45181 standard; protein; 609 AA.
AC ADS45181;
XX
XX 02-DEC-2004 (first entry)
DT Bacterial polypeptide #23611.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX US2003233675-A1.
PN
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 23611; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 609 AA;
Query Match 99.3%; Score 3068; DB 8; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-263;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MCGIVGAIQAQRDVAEILLEGRLRLRYGYDAGLAVDTEGHMTRRLRLKGVQMLAQAAE 60
DB 1 MCGIVGAIQAQRDVAEILLEGRLRLRYGYDAGLAVDTEGHMTRRLRLKGVQMLAQAAE 60
QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHPELREELKARGYTFV 120
DB 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHPELREELKARGYTFV 120
QY 121 SETDTEVAHLVNNELKOGGTILREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSLV 180
DB 121 SETDTEVAHLVNNELKOGGTILREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSLV 180
QY 181 IGLGNGENFIASDQALALPVTRRRFIFLEGGIDAEITRRSVNIFDKTGAEVKRDIESNLQ 240
DB 181 IGLGNGENFIASDQALALPVTRRRFIFLEGGIDAEITRRSVNIFDKTGAEVKRDIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMQKEIYEQNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300
QY 301 CGTNSGMSVRYMFESLAGIPCDVEIASFEFRYKSAVRNSLMTITLSQSGETADTTLA 360
DB 301 CGTNSGMSVRYMFESLAGIPCDVEIASFEFRYKSAVRNSLMTITLSQSGETADTTLA 360
QY 361 RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFALFUSRGDQYPIA 480
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFALFUSRGDQYPIA 480
QY 481 LEGALKUKEISYIHAEAYAGELKHGPLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
DB 481 LEGALKUKEISYIHAEAYAGELKHGPLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGVSSDNMHIIEMPHVEEVIAPFYTPVQLLAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQDAGVSSDNMHIIEMPHVEEVIAPFYTPVQLLAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609
RESULT 13
ADU00590
ID ADU00590 standard; protein; 609 AA.
XX AC ADU00590;
XX 13-JAN-2005 (first entry)
XX Amino acid sequence of a glucosamine-6-phosphate synthase.
XX chitin; chitosan; fermentation;
KW glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;
KW glucosamine-6-phosphate synthase; GNA1 gene; glms gene;
KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;
KW chitin deacetylase; CDA1 gene; CDA2 gene;
KW N-acetylglucosamine-6-phosphate deacetylase;
KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;
KW enzyme.
XX Escherichia coli.
XX WO2004092391-A2.
XX 28-OCT-2004.
XX 12-APR-2004; 2004WO-US011286.
XX 11-APR-2003; 2003US-0462087P.
XX (ARKI-) ARKION LIFE SCI LLC.
XX Deng M, McMullin TW, Grund AD;
XX WPI; 2004-766880/75.
XX N-PSDB; ADU00589.
XX Producing chitin or chitosan, involves culturing microorganism with
PT genetic modification that results in increase in activity of glutamine-
PT fructose-6-phosphate amidotransferase, in fermentation medium, to produce
PT chitin or chitosan.
XX Example 1; SEQ ID NO 23; 161pp; English.
XX The specification describes a method for producing chitin or chitosan by
CC a fermentation process. The method involves culturing in a fermentation
CC medium a microorganism which comprises one or more genetic modifications
CC that result in an increase in the activity of glutamine-fructose-6-
CC phosphate amidotransferase (also known as glucosamine synthetase and
CC glucosamine-6-phosphate synthase, and encoded by the GNA1 eukaryotic gene
CC and the glms bacterial gene), glucosamine-6-phosphate acetyltransferase
CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1
CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-
CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and
CC chitosanase, and collecting the chitin or chitosan. The method is useful
CC for producing chitin or chitosan by utilizing microorganisms such as

CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and
 CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.
 CC cerevisiae, A. niger or A. nidulans. The method enables high quantities
 CC of chitin and chitosan to be produced cost effectively. The present
 CC sequence represents an Escherichia coli glucosamine-6-phosphate synthase.
 CC It was used to transform yeast for use in the method of the invention.
 XX
 XX

SQ Sequence 609 AA;

Query Match	99.3%;	Score 3068;	DB 8;	Length 609;
Best Local Similarity	99.5%;	Pred. No. 2.2e-269;		
Matches 606;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	MCIGVGAIAQRDVAEITLLSGLRLRYGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAE	60
Db	1	MCIGVGAIAQRDVAEITLLSGLRLRYGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAE	60
QY	61	EHLPGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTV	120
Db	61	EHLPGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTV	120
QY	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV	180
Db	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV	180
QY	181	IGLWGENFIASDQALLPVTFRFIFLEGGDIAETIRRSVNIPTDGTGAETKQDIESNLQ	240
Db	181	IGLWGENFIASDQALLPVTFRFIFLEGGDIAETIRRSVNIPTDGTGAETKQDIESNLQ	240
QY	241	YDAGDKGIYCHYMKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYMKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA	300
QY	301	CCTSNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGL	360
Db	301	CCTSNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGL	360
QY	361	RLSKELGYSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLVLLMLVAKL	420
Db	361	RLSKELGYSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLVLLMLVAKL	420
QY	421	SRLLKGLDASIEHDIHVLGQALPSRIEQLMSQDKRIEALAEDEFSDKHALLFLSRGDOYPTA	480
Db	421	SRLLKGLDASIEHDIHVLGQALPSRIEQLMSQDKRIEALAEDEFSDKHALLFLSRGDOYPTA	480
QY	481	LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKKNIEVRAR	540
Db	481	LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKKNIEVRAR	540
QY	541	GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTVDQPR	600
Db	541	GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTVDQPR	600
QY	601	NLAQSVTVE 609	
Db	601	NLAQSVTVE 609	

RESULT 14

ID ADW23840 standard; protein; 608 AA.

XX

AC ADW23840;

XX

DT 07-APR-2005 (first entry)

XX

DE Novel human GFAT protein-related E coli GFAT protein SeqID13.

XX

KW protein purification; glutamine:fructose-6-phosphate amidotransferase;

KW GFAT; antidiabetic; anorectic; metabolic; cytosolic; osteopathic;

KW fungicide; herbicide; diabetes; non-insulin dependent diabetes; obesity;

KW acidosis; cancer; osteoporosis.

XX

OS Escherichia coli.
 XX FR2857374-Al.
 XX 14-JAN-2005.
 XX 08-JUL-2003; 2003FR-00008350.
 XX 08-JUL-2003; 2003FR-00008350.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Badet DMAJE, Badet BF;
 XX WPI; 2005-076419/09.
 XX New modified glutamine:fructose-6-phosphate amidotransferase, useful for
 PT identifying specific inhibitors, potentially useful for treating e.g.
 PT diabetes, contains a purification tag, also related nucleic acid.
 XX Disclosure; SEQ ID NO 13; 61pp; French.
 XX This invention relates to a novel protein corresponding to human
 CC enzymatically active but modified glutamine:fructose-6-phosphate
 CC amidotransferase (GFAT). The invention may be useful for the production
 CC of compounds with an antidiabetic, anorectic, metabolic, cytosolic,
 CC osteopathic, fungicide or herbicide activity. The invention may be used
 CC to screen for agents that modify, especially inhibit, its activity,
 CC (particularly type II), obesity, acidosis, ketosis, cancer and
 CC osteoporosis. In addition inhibitors of plant and fungal GFAT may be
 CC useful as herbicides and fungicides, respectively. The present sequence
 CC is that of a human GFAT-like protein of the invention.

SQ Sequence 608 AA;

Query Match	99.2%;	Score 3063;	DB 9;	Length 608;
Best Local Similarity	99.5%;	Pred. No. 6.3e-269;		
Matches 605;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	2	CGIVGAIAQRDVAEITLLSGLRLRYGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAE	61
Db	1	CGIVGAIAQRDVAEITLLSGLRLRYGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAE	60
QY	62	HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTVFS	121
Db	61	HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTVFS	120
QY	122	ETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLVI	181
Db	121	ETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLVI	180
QY	182	GLMGGENFIASDQALLPVTFRFIFLEGGDIAETIRRSVNIPTDGTGAETKQDIESNLQY	241
Db	181	GLMGGENFIASDQALLPVTFRFIFLEGGDIAETIRRSVNIPTDGTGAETKQDIESNLQY	240
QY	242	DAGDKGIYCHYMKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILAC	301
Db	241	DAGDKGIYCHYMKEIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILAC	300
QY	302	GTSYNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGLR	361
Db	301	GTSYNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGLR	360
QY	362	LSKELGYSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLVLLMLVAKLS	421
Db	361	LSKELGYSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLVLLMLVAKLS	420
QY	422	RUKGLDASIEHDIHVLGQALPSRIEQLMSQDKRIEALAEDEFSDKHALLFLSRGDOYPIAL	481
Db	421	RUKGLDASIEHDIHVLGQALPSRIEQLMSQDKRIEALAEDEFSDKHALLFLSRGDOYPIAL	480
QY	482	EGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKKNIEVRARG	541

Db	481	EGALKLKEISYTHAEAYAGELKKGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRARG	540
Qy	542	GQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPRN	601
Db	541	GQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPRN	600
Qy	602	LAKSVTVE 609	
Db	601	LAKSVTVE 608	
RESULT 15			
AAV58825			
ID	AAV58825 standard; protein; 609 AA.		
XX			
AC	AAV58825;		
XX			
DT	08-MAY-2000 (first entry)		
XX			
DE	E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-124.		
XX			
KW	Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-124;		
KW	glucosamine; metabolic engineering; plasmid pKLN23-124; mutein.		
XX			
OS	Escherichia coli.		
XX			
FH	Key Location/Qualifiers		
FT	Misc-difference 469		
FT	/note= "replaces wild-type Leu"		
XX			
PN	W0200004182-A1.		
XX			
PD	27-JAN-2000.		
XX			
PP	15-JUL-1999; 99WO-US015976.		
XX			
PR	15-JUL-1998; 98US-00115475.		
XX			
PA	(DCVB-) DCV INC DBA BIO-TECH RESOURCES.		
XX			
PI	Berry A, Burlingame RP, Millis JR;		
XX			
DR	WPI; 2000-182441/16.		
DR	N-PSDB; AAZ58252.		
XX			
PT	Fermentation of E. coli having an altered amino acid sugar metabolic		
PT	pathway to produce glucosamine, especially using novel recombinant		
PT	variant glucosamine-6-phosphate synthases.		
XX			
PS	Claim 28; Page 130-132; 150pp; English.		
XX			
CC	The present sequence is that of a mutant, denoted GlcN6P-S-124, of the		
CC	glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.		
CC	When compared with the wild-type sequence (see AAY58822), the mutant		
CC	includes a Leu-469 to Thr amino acid substitution. This alteration is		
CC	predicted from the mutated glms gene in plasmid pKLN23-124 (see AAZ58252).		
CC	The invention provides methods for the overproduction of glucosamine by		
CC	fermentation using a genetically engineered microorganism, especially E.		
CC	coli, that includes a modified GlcN6P synthase. Production of the		
CC	glucosamine by recombinant strain 2123-124(pKLN23-124) was significantly		
CC	increased when compared to a strain expressing wild-type Glc6NP synthase		
CC	owing to reduced product inhibition		
XX			
SQ	Sequence 609 AA;		
Query Match 99.1%; Score 3061; DB 3; Length 609;			
Best Local Similarity 99.3%; Pred. No. 9.7e-269;			
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVQMLAQAAE	60
Db	1	MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVQMLAQAAE	60

Qy	61	EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNHEPLREELKARGYTFV	120
Db	61	EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNHEPLREELKARGYTFV	120
Qy	121	SETDTEVIAHLVNMELKOGGTLEAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNMELKOGGTLEAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV	180
Qy	181	IGLGMGENFIASDQALLPVTRRFIFLEEGDIAIITRRSVNIFDKTGAEVKRODIESNLQ	240
Db	181	IGLGMGENFIASDQALLPVTRRFIFLEEGDIAIITRRSVNIFDKTGAEVKRODIESNLQ	240
Qy	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA	300
Qy	301	CGTSYNSGMVSRYPFESLAGIPCDVEIASBFRYKSAVRRNSLMTLSQSGETADTLA	360
Db	301	CGTSYNSGMVSRYPFESLAGIPCDVEIASBFRYKSAVRRNSLMTLSQSGETADTLA	360
Qy	361	RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL	420
Db	361	RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL	420
Qy	421	SRKGLDASIEHDIVHGLQALPSRIEOMLSQDKRIEALAEDFSXKHHALFLSRGDQYPIA	480
Db	421	SRKGLDASIEHDIVHGLQALPSRIEOMLSQDKRIEALAEDFSXKHHALFLSRGDQYPIA	480
Qy	481	LEGALKKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR	540
Db	481	LEGALKKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR	540
Qy	541	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR	600
Db	541	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	

Search completed: June 14, 2006, 15:21:11
Job time : 290.762 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM protein - protein search, using sw model

Run on: June 14, 2006, 15:28:50 ; Search time 209.344 Seconds
(without alignments)
1347.534 Million cell updates/sec

Title: US-10-612-779-6

Perfect score: 3089

Sequence: 1 MCGIVCAIAQRDVAEILLEG.....LIKGTVDQPRNAKSTVTE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3089	100.0	609	4	US-10-024-460-22
2	3089	100.0	609	4	US-10-612-779-6
3	3089	100.0	609	5	US-10-823-397-25
4	3072	99.4	609	4	US-10-024-460-28
5	3072	99.4	609	4	US-10-024-460-31
6	3072	99.4	609	4	US-10-612-779-10
7	3072	99.4	609	4	US-10-612-779-12
8	3068	99.3	609	3	US-09-815-242-10399
9	3068	99.3	609	4	US-10-024-460-16
10	3068	99.3	609	4	US-10-369-493-23611
11	3068	99.3	609	4	US-10-282-122A-56753
12	3068	99.3	609	4	US-10-612-779-2
13	3068	99.3	609	5	US-10-823-397-23
14	3061	99.1	609	4	US-10-024-460-25
15	3061	99.1	609	4	US-10-612-779-8
16	3053	98.8	609	4	US-10-024-460-19
17	3053	98.8	609	4	US-10-612-779-4
18	3043	98.5	609	4	US-10-612-779-14
19	3035	98.3	609	3	US-09-815-242-14065
20	3035	98.3	609	4	US-10-282-122A-76152
21	2932	94.9	609	4	US-10-282-122A-56279
22	2862	92.7	609	3	US-09-815-242-11758
23	2857	92.5	609	4	US-10-282-122A-59335
24	2700.5	87.4	608	4	US-10-282-122A-72965
25	2668	86.4	609	4	US-10-282-122A-78533
26	2647	85.7	609	4	US-10-369-493-21242
27	2605	84.3	609	4	US-10-369-493-438

ALIGNMENTS

RESULT 1

US-10-024-460-22

; Sequence 22, Application US/10024460

; Publication No. US20030044939A1

; GENERAL INFORMATION:

; APPLICANT: Berry, Alan

; APPLICANT: Burlingame, Richard P.

; APPLICANT: Millis, James R.

; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE

; FILE REFERENCE: 3161-18-C1

; CURRENT APPLICATION NUMBER: US/10/024,460

; CURRENT FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-10-024-460-22

Query Match 100.0%; Score 3089; DB 4; Length 609;

Best Local Similarity 100.0%; Pred. No. 4.2e-252;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVCAIAQRDVAEILLEGRLRLEVRGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAAE 60

Db 1 MCGIVCAIAQRDVAEILLEGRLRLEVRGVDSAGLAVVDTEGHMTRRLRLKGVQMLAAQAAE 60

QY 61 EPHLHGCTGIAITRWATHGEPSEVNAHPVSHIIVVHNGIIEHHEPLREELKARGYTFV 120

Db 61 EPHLHGCTGIAITRWATHGEPSEVNAHPVSHIIVVHNGIIEHHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180

Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180

QY 181 IGLGMGENTIASDQALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240

Db 181 IGLGMGENTIASDQALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240

QY 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIOILA 300

Db 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIOILA 300

QY 301 CGTSSNGSMVSRVWFESLAGIPCDVFIASEFRYKSAVRNRSLMITLSQSGETADTLAGL 360

Db 301 CGTSSNGSMVSRVWFESLAGIPCDVFIASEFRYKSAVRNRSLMITLSQSGETADTLAGL 360

Sequence 69028, A
Sequence 11018, A
Sequence 58163, A
Sequence 313, App
Sequence 77063, A
Sequence 67388, A
Sequence 67654, A
Sequence 13847, A
Sequence 12111, A
Sequence 6848, A
Sequence 69638, A
Sequence 44943, A
Sequence 15424, A
Sequence 15792, A
Sequence 16174, A
Sequence 61472, A
Sequence 19598, A
Sequence 9210, Ap

```
Db 301 COTSNSGMVSRYPFESLAGIPCDVEIASFERYKSAVRNSLMTLSQSGETADTLAGL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Qy 601 NLAKSVTVVE 609
Db 601 NLAKSVTVVE 609

RESULT 2
US-10-612-779-6
; Sequence 6, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Processes and Materials for Production of Glucosamine and N-Acetyls
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612, 779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393, 348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-6

Query Match 100.0%; Score 3089; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-252;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKQVQLAAQAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKQVQLAAQAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPRLREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPRLREELKARGYTFV 120
Qy 121 SETDTEVIHLVNNELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIHLVNNELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
```

```
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYCHYMKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Qy 301 COTSNSGMVSRYPFESLAGIPCDVEIASFERYKSAVRNSLMTLSQSGETADTLAGL 360
Db 301 COTSNSGMVSRYPFESLAGIPCDVEIASFERYKSAVRNSLMTLSQSGETADTLAGL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Qy 601 NLAKSVTVVE 609
Db 601 NLAKSVTVVE 609

RESULT 3
US-10-823-397-25
; Sequence 25, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823, 397
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462, 087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-823-397-25

Query Match 100.0%; Score 3089; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-252;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKQVQLAAQAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLGKQVQLAAQAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPRLREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPRLREELKARGYTFV 120
Qy 121 SETDTEVIHLVNNELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIHLVNNELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKIIEYQPNNAIKNTLTGRISHGOVDLSSELGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKIIEYQPNNAIKNTLTGRISHGOVDLSSELGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVYFESLAGIPCDVEIASEFRYKRSVAVRRNSLMITLSQSGETADTLA 360
DB 301 CGTSYNSGMVSRVYFESLAGIPCDVEIASEFRYKRSVAVRRNSLMITLSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQYPIA 480
DB 421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVPADODAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVQDPR 600
DB 541 GGOLYVPADODAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAQSVTVE 609
DB 601 NLAQSVTVE 609

RESULT 4
US-10-024-460-28
; Sequence 28, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-28

Query Match 99.4%; Score 3072; DB 4; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.1e-250;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
DB 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
QY 61 EHPHLCGTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLRELKARGYTFV 120
DB 61 EHPHLCGTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLRELKARGYTFV 120
QY 121 SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSSPLV 180
DB 121 SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKIIEYQPNNAIKNTLTGRISHGOVDLSSELGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKIIEYQPNNAIKNTLTGRISHGOVDLSSELGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVYFESLAGIPCDVEIASEFRYKRSVAVRRNSLMITLSQSGETADTLA 360
DB 301 CGTSYNSGMVSRVYFESLAGIPCDVEIASEFRYKRSVAVRRNSLMITLSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQYPIA 480
DB 421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVPADODAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVQDPR 600
DB 541 GGOLYVPADODAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAQSVTVE 609
DB 601 NLAQSVTVE 609

RESULT 5
US-10-024-460-31
; Sequence 31, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-31

Query Match 99.4%; Score 3072; DB 4; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.1e-250;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
DB 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
QY 61 EHPHLCGTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLRELKARGYTFV 120
DB 61 EHPHLCGTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLRELKARGYTFV 120
QY 121 SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSSPLV 180
DB 121 SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
```

241	Qy	YDAGKGIIYCHYMOKEIIYEQPNAINKTUVTGRISHGQVDLSLGPNADELLSKVEHIQILA	300
241	Db	YDAGDKGIYRYHMOKEIIYEQPNAINKTUVTGRISHGQVDLSLGPNADELLSKVEHIQILA	300
301	Qy	CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFYRKSAVRNLSMTLSOSGETADTLAGL	360
301	Db	CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFYRKSAVRNLSMTLSOSGETADTLAGL	360
361	Qy	RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTQTLTVLLMLVAKL	420
361	Db	RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTQTLTVLLMLVAKL	420
421	Qy	SRLKGLDASIEHDIYVHGLOALPSRIEQMLSQDKRIEALAEFSDKHHALFLSRGQOYPTA	480
421	Db	SRLKGLDASIEHDIYVHGLOALPSRIEQMLSQDKRIEALAEFSDKHHALFLSRGQOYPTA	480
481	Qy	LEGALKLKEISYIHAEAYAAGELKHGPLALTDADMPVIWVAPNNLELKLKSNIEVRAR	540
481	Db	LEGALKLKEISYIHAEAYAAGELKHGPLALTDADMPVIWVAPNNLELKLKSNIEVRAR	540
541	Qy	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDQPR	600
541	Db	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTDVDQPR	600
601	Qy	NLAKSVTVE	609
601	Db	NLAKSVTVE	609

RESULT 6
 US-10-612-779-10
 ; Sequence 10, Application US/10612779
 ; Publication No. US20040091976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deng, Ming-De
 ; APPLICANT: Angerer, J. David
 ; APPLICANT: Cyron, Don
 ; APPLICANT: Grund, Alan
 ; APPLICANT: Jerrell Jr., Thomas
 ; APPLICANT: Leanna, Candice
 ; APPLICANT: Mathre, Owen
 ; APPLICANT: Rosson, Reinhardt
 ; APPLICANT: Running, Jeff
 ; APPLICANT: Severson, Dave
 ; APPLICANT: Song, Linsheng
 ; APPLICANT: Wassink, Sarah
 ; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetylg
 ; TITLE OF INVENTION: Glucosamine
 ; FILE REFERENCE: 3161-18-2
 ; CURRENT APPLICATION NUMBER: US/10/612,779
 ; CURRENT FILING DATE: 2003-07-01
 ; PRIOR APPLICATION NUMBER: 60/393,348
 ; PRIOR FILING DATE: 2002-07-01
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-612-779-10

[illegible]

```

RESULT 7
US-10-612-779-12
; Sequence 12, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetylg
; TITLE OF INVENTION: Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-612-779-12

```

Qy	61	EHPHGGTGIAHTRWATHG	BSP	EVNAHPVSEHI	VVHNGII	ENHEPLREELKARGYTFV	120
D6	61	EHPHGGTGIAHTRWATHG	BSP	EVNAHPVSEHI	VVHNGII	ENHEPLREELKARGYTFV	120

```
QY 1 MCGIVGAI AQRDVAEILLEGRLRLRYGDSAGLAVVDTEGHMTRRLRLKVKVQLAAAE 60
Db 1 MCGIVGAI AQRDVAEILLEGRLRLRYGDSAGLAVVDTEGHMTRRLRLKVKVQLAAAE 60
QY 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHYSEHIVVHNGIIEHNEHPLREELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHYSEHIVVHNGIIEHNEHPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGCTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGCTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRI SHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRI SHGQVDLSSELGNADDELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGQDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGQDQYPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

RESULT 8

US-09-815-242-10399

; Sequence 10399, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITPA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10399

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-815-242-10399

Query Match 99.3%; Score 3068; DB 3; Length 609;

Best Local Similarity 99.5%; Pred. No. 2.5e-250; Indels 0; Gaps 0;

Matches 606; Conservative 0; Mismatches 35;

```
QY 1 MCGIVGAI AQRDVAEILLEGRLRLRYGDSAGLAVVDTEGHMTRRLRLKVKVQLAAAE 60
Db 1 MCGIVGAI AQRDVAEILLEGRLRLRYGDSAGLAVVDTEGHMTRRLRLKVKVQLAAAE 60
QY 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHYSEHIVVHNGIIEHNEHPLREELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHYSEHIVVHNGIIEHNEHPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGCTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGCTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRI SHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRI SHGQVDLSSELGNADDELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGQDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGQDQYPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

RESULT 9

US-10-024-460-16

; Sequence 16, Application US/10024460

; Publication No. US20030044939A1

; GENERAL INFORMATION:

; APPLICANT: Berry, Alan

; APPLICANT: Burlingame, Richard P.

; APPLICANT: Millis, James R.

; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE

; FILE REFERENCE: 3161-18-C1

; CURRENT APPLICATION NUMBER: US/10/024,460


```
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-16

Query Match      99.3%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.5e-250;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKQVQLAAQAAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKQVQLAAQAAE 60
QY 61 EHPHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLVFAFDQDAGFVSSDNMHIEMPHVEEVIAPITYVPLQLLAYHVHVALIKGTVDQDPR 600
Db 541 GGQLVFAFDQDAGFVSSDNMHIEMPHVEEVIAPITYVPLQLLAYHVHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 10
US-10-369-493-23611
; Sequence 23611, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23611
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23611

Query Match      99.3%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.5e-250;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKQVQLAAQAAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKQVQLAAQAAE 60
QY 61 EHPHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLVFAFDQDAGFVSSDNMHIEMPHVEEVIAPITYVPLQLLAYHVHVALIKGTVDQDPR 600
Db 541 GGQLVFAFDQDAGFVSSDNMHIEMPHVEEVIAPITYVPLQLLAYHVHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
US-10-282-122A-56753
; Sequence 56753, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```


; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 56753
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-282-122A-56753

 Query Match 99.3%; Score 3068; DB 4; Length 609;
 Best Local Similarity 99.5%; Pred. No. 2.5e-250;
 Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 QY 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60
 DB 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60

 QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
 DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

 QY 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
 DB 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180

 QY 181 IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTCAGAEVKRODIESNLQ 240
 DB 181 IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTCAGAEVKRODIESNLQ 240

 QY 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRISHGOVDLSELGNADLLSKVEHIQILA 300
 DB 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRISHGOVDLSELGNADLLSKVEHIQILA 300

 QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQSGETADTLAAGL 360
 DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQSGETADTLAAGL 360

 QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTQTLTLLMLVAKL 420
 DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTQTLTLLMLVAKL 420

 QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEFSDKHHALFLRGDQVPIA 480

Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEFSDKHHALFLRGDQVPIA 480
 QY 481 LEGALKLKEISYTHAEAYAGELKHGCPALIIDADMPVIVVAPNNELEKLSNIEVRAR 540
 Db 481 LEGALKLKEISYTHAEAYAGELKHGCPALIIDADMPVIVVAPNNELEKLSNIEVRAR 540
 QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDDVQPR 600
 Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDDVQPR 600
 QY 601 NLAKSVTVE 609
 Db 601 NLAKSVTVE 609

 RESULT 12
 US-10-612-779-2
 ; Sequence 2, Application US/10612779
 ; Publication No. US20040091976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deng, Ming-De
 ; APPLICANT: Angerer, J. David
 ; APPLICANT: Cyron, Don
 ; APPLICANT: Grund, Alan
 ; APPLICANT: Jerrell Jr., Thomas
 ; APPLICANT: Leanna, Candice
 ; APPLICANT: Mathre, Owen
 ; APPLICANT: Rosson, Reinhardt
 ; APPLICANT: Running, Jeff
 ; APPLICANT: Severson, Dave
 ; APPLICANT: Song, Linsheng
 ; APPLICANT: Wassink, Sarah
 ; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
 ; TITLE OF INVENTION: Glucosamine
 ; FILE REFERENCE: 3161-18-2
 ; CURRENT APPLICATION NUMBER: US/10/612,779
 ; CURRENT FILING DATE: 2003-07-01
 ; PRIOR APPLICATION NUMBER: 60/393,348
 ; PRIOR FILING DATE: 2002-07-01
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-612-779-2

 Query Match 99.3%; Score 3068; DB 4; Length 609;
 Best Local Similarity 99.5%; Pred. No. 2.5e-250;
 Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 QY 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60
 Db 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60

 QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
 Db 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

 QY 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
 Db 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180

 QY 181 IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTCAGAEVKRODIESNLQ 240
 Db 181 IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTCAGAEVKRODIESNLQ 240

 QY 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRISHGOVDLSELGNADLLSKVEHIQILA 300
 Db 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRISHGOVDLSELGNADLLSKVEHIQILA 300

 QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQSGETADTLAAGL 360

Db 301 CGTSYNSGMVSRYPFESLAGICDVEIASSEPRYKSAVRNLSMITLSQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 13
US-10-823-397-23
; Sequence 23, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823,397
; PRIOR FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-823-397-23

Query Match 99.3%; Score 3068; DB 5; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.5e-250;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRRLKGVQMLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRRLKGVQMLAQAAE 60
QY 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDQIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDQIESNLQ 240
QY 241 YDAGDKGYCHYMOKEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLELLSKVEHIQILA 300
Db 241 YDAGDKGYCHYMOKEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGICDVEIASSEPRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGICDVEIASSEPRYKSAVRNLSMITLSQSGETADTLA 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 14
US-10-024-460-25
; Sequence 25, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-25

Query Match 99.1%; Score 3061; DB 4; Length 609;
Best Local Similarity 99.3%; Pred. No. 9.7e-250;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRRLKGVQMLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRRLKGVQMLAQAAE 60
QY 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDQIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDQIESNLQ 240
QY 241 YDAGDKGYCHYMOKEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLELLSKVEHIQILA 300
Db 241 YDAGDKGYCHYMOKEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGICDVEIASSEPRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGICDVEIASSEPRYKSAVRNLSMITLSQSGETADTLA 360

```
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLSRGDQVPIA 480
DB 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHAPPLGRGDQVPIA 480
QY 481 LEGALKLKEISYTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
DB 541 GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 15
US-10-612-779-8
; Sequence 8, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Rosson, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; TITLE OF INVENTION: Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612, 779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-8

Query Match 99.1%; Score 3061; DB 4; Length 609;
Best Local Similarity 99.3%; Pred. No. 9.7e-250;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGIVGAIARDVAEILLBGLRRLRYGYDSAGLAVVDTEGHMTRRLRLGKQVQLAAAE 60
DB 1 MCGIVGAIARDVAEILLBGLRRLRYGYDSAGLAVVDTEGHMTRRLRLGKQVQLAAAE 60
QY 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIENHEPLREELKARGYTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSEHIVVHNGIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV 180
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV 180
QY 181 IGLMGENTIASDQALLPVTRRFIFLEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240
DB 181 IGLMGENTIASDQALLPVTRRFIFLEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADELLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVYWFESLAGIPCDVEIASEFRYRKSARRNSLMTILSQSGETADTLAGL 360
DB 301 CGTSYNSGMVSRVYWFESLAGIPCDVEIASEFRYRKSARRNSLMTILSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLSRGDQVPIA 480
DB 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHAPPLGRGDQVPIA 480
QY 481 LEGALKLKEISYTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
DB 541 GGOLYVFADODAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609
```

Search completed: June 14, 2006, 15:34:24
Job time : 211.344 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: June 14, 2006, 15:28:50 ; Search time 54.6562 Seconds
 (without alignments)
 1347.534 Million cell updates/sec
 Title: US-10-612-779-30
 Perfect score: 835
 Sequence: 1 MSLPDGFYIRMEEGDLEQV.....VKFYKCGFSNAGVEMQIRK 159
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA Main:
 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	100.0	159	4	US-10-429-812-4
2	835	100.0	159	4	US-10-612-779-30
3	835	100.0	159	5	US-10-823-397-33
4	333	39.9	149	4	US-10-032-585-7301
5	328	39.3	149	4	US-10-612-779-32
6	328	39.3	149	5	US-10-823-397-35
7	275	32.9	190	4	US-10-128-714-8300
8	265.5	31.8	177	4	US-10-425-115-236099
9	261.5	31.3	149	4	US-10-612-779-34
10	234.5	28.1	174	4	US-10-424-599-159788
11	216.5	25.9	163	4	US-10-425-115-261686
12	216.5	25.9	174	4	US-10-425-114-61424
13	209	25.0	184	3	US-09-795-926-46
14	209	25.0	184	4	US-10-364-774-46
15	209	25.0	184	6	US-11-134-241-46
16	208	24.9	184	3	US-09-798-029-4
17	207	24.8	219	6	US-11-097-143-1509
18	204	24.4	184	4	US-10-092-900A-26
19	202.5	24.3	180	4	US-10-128-714-3300
20	202	24.2	184	4	US-10-092-900A-28
21	201.5	24.1	166	4	US-10-437-963-136699
22	199.5	23.9	165	4	US-10-437-963-116060
23	142	17.0	98	4	US-10-424-599-252530
24	116.5	14.0	157	6	US-11-045-004-2748
25	116	13.9	160	4	US-10-156-761-12507
26	114	13.7	133	3	US-09-798-029-25
27	112	13.4	140	4	US-10-282-122A-51716

28 110 13.2 179 3 US-09-738-626-4615 Sequence 4615, Ap
 29 110 13.2 179 5 US-10-703-7998-20 Sequence 20, Appl
 30 106 12.7 185 6 US-11-045-004-264 Sequence 264, Appl
 31 105.5 12.6 159 4 US-10-433-256-9 Sequence 9, Appl
 32 101 12.1 168 6 US-10-282-122A-60831 Sequence 60831, A
 33 101 12.1 168 6 US-11-045-004-1735 Sequence 1735, Ap
 34 101 12.1 170 6 US-11-124-367A-459 Sequence 459, Appl
 35 98 11.7 23 3 US-09-927-734C-5 Sequence 5, Appl
 36 97.5 11.7 147 5 US-10-501-282-2468 Sequence 2468, Ap
 37 97.5 11.7 197 5 US-10-501-282-2470 Sequence 2470, Ap
 38 97 11.6 218 4 US-10-425-115-339875 Sequence 339875,
 39 96 11.5 254 4 US-10-437-963-160771 Sequence 160771,
 40 96 11.5 257 4 US-10-424-599-273495 Sequence 273495,
 41 96 11.5 263 4 US-10-425-114-50819 Sequence 50819, A
 42 95 11.4 104 4 US-10-425-115-275860 Sequence 275860,
 43 95 11.4 371 4 US-10-425-115-315609 Sequence 315609,
 44 94.5 11.3 151 6 US-11-045-004-1609 Sequence 1609, Ap
 45 93 11.1 192 4 US-10-424-599-147429 Sequence 147429,

ALIGNMENTS

RESULT 1
 US-10-429-812-4
 ; Sequence 4, Application US/10429812
 ; Publication No. US20040003432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Obukowicz, Mark
 ; TITLE OF INVENTION: THE PRODUCTION OF HEXOSAMINES AND USES THEREFORE
 ; FILE REFERENCE: 18438/09004
 ; CURRENT APPLICATION NUMBER: US/10/429,812
 ; CURRENT FILING DATE: 2003-05-05
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 159
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-429-812-4

Query Match 100.0%; Score 835; DB 4; Length 159;
 Best Local Similarity 100.0%; Pred. No. 1.1e-82;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLPDGFYIRMEEGDLEQVTTKVLTVGTITPESFSLIKYWNATVWVNDKKIM 60
 DB 1 MSLPDGFYIRMEEGDLEQVTTKVLTVGTITPESFSLIKYWNATVWVNDKKIM 60
 QY 61 QYNPWVVDKRTETVAATGNIIEERKIIHELGCCHIEDIAVNSKYOGGKLLIDQLV 120
 DB 61 QYNPWVVDKRTETVAATGNIIEERKIIHELGCCHIEDIAVNSKYOGGKLLIDQLV 120
 QY 121 TIGFDYGCYKIIIDCDCKNKVFEKCGFSNAGVEMQIRK 159
 DB 121 TIGFDYGCYKIIIDCDCKNKVFEKCGFSNAGVEMQIRK 159

RESULT 2
 US-10-612-779-30
 ; Sequence 30, Application US/10612779
 ; Publication No. US20040091976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deng, Ming-De
 ; APPLICANT: Angerer, J. David
 ; APPLICANT: Cyron, Don
 ; APPLICANT: Grund, Alan
 ; APPLICANT: Jerrell Jr., Thomas
 ; APPLICANT: Leanna, Candice
 ; APPLICANT: Mathre, Owen
 ; APPLICANT: Rosson, Reinhardt
 ; APPLICANT: Running, Jeff
 ; APPLICANT: Severson, Dave

```
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-612-779-30

Query Match      100.0%; Score 835; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKIM 60
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKIM 60

Qy 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGLLIDQLV 120
Db 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGLLIDQLV 120

Qy 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
Db 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159

RESULT 3
US-10-823-397-33
; Sequence 33, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823,397
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-823-397-33

Query Match      100.0%; Score 835; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKIM 60
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKIM 60

Qy 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGLLIDQLV 120
Db 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGLLIDQLV 120

Qy 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
Db 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
```

```
RESULT 4
US-10-032-585-7301
; Sequence 7301, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7301
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7301

Query Match      39.9%; Score 333; DB 4; Length 149;
Best Local Similarity 44.0%; Pred. No. 6.2e-28;
Matches 70; Conservative 26; Mismatches 51; Indels 12; Gaps 3;

Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKI 59
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSLIKYWNATVWNNDEKKI 59

Qy 60 MQYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGLLIDQL 119
Db 51 SIYHPYVITN-ASGIVVATGMLFVEKKLIHECGKVGHIEDISVAKSEQGKGLGYLVTS 109

Qy 120 VTIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIR 158
Db 110 TKVAQENDCYKVIILDCSPENVGFYKCGYKDGGMVCR 148
```

```
RESULT 5
US-10-612-779-32
; Sequence 32, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-612-779-32

Query Match      39.3%; Score 328; DB 4; Length 149;
Best Local Similarity 43.4%; Pred. No. 2.2e-27;
```

	Matches	69;	Conservative	26;	Mismatches	52;	Indels	12;	Gaps	3;
Qy	1	MSLPDGFVIRMERGDLLE-QVTETKLVLTTVGTTTPESFSLKIYKYNREATVMNDNEDKKI	59							
Dd	1	MLPQGTFTRKLKTVDYDNQYLETKLVLTTVGEISKEDEFTELYNHWSLSP-----	50							
Qy	60	MQYNPMVIVDKRTTVAATGNIIRKIIHIEGLCGHIEDIAVNSKTKQGOGGLKLLIDQL	119							
Dd	51	SIYHPYVITN-ASGIIVATGMFLVEKLLIHCEGVGHIEDISVAKSEGGKLGYYLVTSL	109							
Qy	120	VTIGFDYGCVYKIILDCDEKNVKVFYEKCGFSNAGVEMQIR	158							
Dd	110	TKVAOENDCYKVIDLCSPENYNGVFYEKCGYKDGGYEMVCR	148							

RESULT 6
US-10-823-397-35
Sequence 35, Application US/10823397
Publication No. US20050042735A1
GENERAL INFORMATION:
APPLICANT: McMullin, Thomas
APPLICANT: Ding, Ming-De
APPLICANT: Grund, Alan
TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
TITLE OF INVENTION: Chitosan in Microorganisms
FILE REFERENCE: 3161-18-3
CURRENT APPLICATION NUMBER: US/10/823,397
CURRENT FILING DATE: 2004-04-12
PRIOR APPLICATION NUMBER: 60/462,087
PRIOR FILING DATE: 2003-04-11
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn version 3.2
SEQ ID NO 35
LENGTH: 149
TYPE: PRT
ORGANISM: Candida albicans
US-10-823-397-35

[illegible]

```

RESULT 7
US-10-128-714-8300
; Sequence 8300, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshekin, Alexey M
; APPLICANT: Lemieux, Sebastian M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23

```

```

; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8300
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8300

```

Query Match	32.9%;	Score	275;	DB	4;	Length	190;	
Best Local Similarity	41.3%;	Pred.	No. 1.8e-21;					
Matches	64;	Conservative	19;	Mismatches	55;	Indels	14; Gaps	5;
<hr/>								
Qy	3	LPDGFYRRMBEGGLEQ-VTETLKVLTVTGVIITPESFKLIKYNEATVW-NDNEDKKIM	60					
Dd	44	LPADYTRPLCRSDYKKGYLVDLVLTVTGVGINEEQ-----WNRSRYEWIARSD----	92					
<hr/>								
Qy	61	QYNPMWIVDKRTETVAATGNIIIBRKIIHELGLCHIEDIAVNSKYQQGGIGKLIDIQIV	120					
Dd	93	EYYLLVVCDEGR-IVGTSVLVERKPIHSLGMVGHIEDIAVEKQQQKKGLRIIQALD	151					
<hr/>								
Qy	121	TIGFDYGCIKIILDCDEKNVKFYKCGFSNAGVEM	155					
<hr/>								
Dd	152	YVASKGCYCYTIIDCSBANEGFYTKCFPKRAGLEM	186					

```

RESULT 8
US-10-425-115-236099
; Sequence 236099, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 236099
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(177)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_146905C.1.pep
US-10-425-115-236099

```

	Query Match	31.8%	Score 265.5;	DB 4;	Length 177;
	Best Local Similarity	41.1%;	Pred. No. 1.8e-20;		
	Matches	65;	Conservative 21;	Mismatches 57;	Indels 15; Gaps 5
Qy	2	SLPDGFYRMBEGLDLEQ-VTETLKVLTIVGTITPESPSKLIKYNNEATVWNDNEDKKIM	60		
	:	:::::	:	:	:
Dd	19	ALPEGYTLRALRKSDFNNGFLDCLRVLTIVGDTITEADFYK-QYDDMAAGS-----	68		
	:	:::::	:	:	:
Qy	61	QYNPMWVD--KRTEVAATGNIIIERKIIHELGCGHIE-DIAVNSKYQGOGKLGLIID	117		
	:	:::::	:	:	:
Dd	69	-YYIITIDTSRKNPNPVGTGALTITERFIHSLGAVGHIEXDIAVKDQQGKLGLRIIQ	127		
	:	:::::	:	:	:
Qy	118	QLVATIGFDYGCVKIILDCDEKNRVFYEKCQGSNAGVM	155		
	:	:::::	:	:	:


```
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61424
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-059-G11_FLI.pep
US-10-425-114-61424

Query Match      25.9%; Score 216.5; DB 4; Length 174;
Best Local Similarity 33.8%; Pred. No. 3.9e-15;
Matches 52; Conservative 28; Mismatches 65; Indels 9; Gaps 3;

Qy 5 DGFYIRMEGDLQ-VTETLKVLTVTGTTIPESFSLIKYVNEATVWVNDNEDKKIMQYN 63
Db 27 DTIHRLECSHERGFVALLSQLSPCLDTSVPATRP-----AELAAQGDHIIILVAE 81

Qy 64 PMWIVDKRETVAATGNIIIRKIIHELGLCHIEDIAVNSKYQGGLGKLLIDOLVTVIG 123
Db 82 DPSASDRR---ILATGCLFVERKFKRGCGKHVEDVVVDAARGLGLRIVRLVVIS 138

Qy 124 FDYGCYKIILDCDEKNVIFYEKCQFSNAGVEMQI 157
Db 139 RDAGCYKVIILDCETPELAYAYAKCGFVEKGVQNAV 172

RESULT 13
US-09-795-926-46
; Sequence 46, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-795-926-46

Query Match      25.0%; Score 209; DB 3; Length 184;
Best Local Similarity 32.9%; Pred. No. 2.8e-14;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEGDLQ-VTETLKVLTVTGTTIPESFSLIKYVNEATVWVNDNEDKKIMQYN 63
Db 37 EGLVLRPLCTADLNRGFFKVLQLTETGVVSPQFMKSFHM-----KXSGDY 85

Qy 64 PMWIVDKRETVAATGNIIIRKIIHELGLCHIEDIAVNSKYQGGLGKLLIDOLVTVIG 123
Db 86 VTVVEDVTLGQIVATATLIIEHKFIHSCAKRGVEDVVVSDCECRGKQLGKLLSTLTLS 145

Qy 124 FDYGCYKIILDCDEKNVIFYEKCQFS 149
Db 146 KKLNCYKITLECLPQNVGFYKKVGYT 171

RESULT 14
US-10-364-774-46
; Sequence 46, Application US/10364774
; Publication No. US2003014497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-364-774-46

Query Match      25.0%; Score 209; DB 4; Length 184;
Best Local Similarity 32.9%; Pred. No. 2.8e-14;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEGDLQ-VTETLKVLTVTGTTIPESFSLIKYVNEATVWVNDNEDKKIMQYN 63
Db 37 EGLVLRPLCTADLNRGFFKVLQLTETGVVSPQFMKSFHM-----KXSGDY 85

Qy 64 PMWIVDKRETVAATGNIIIRKIIHELGLCHIEDIAVNSKYQGGLGKLLIDOLVTVIG 123
Db 86 VTVVEDVTLGQIVATATLIIEHKFIHSCAKRGVEDVVVSDCECRGKQLGKLLSTLTLS 145

Qy 124 FDYGCYKIILDCDEKNVIFYEKCQFS 149
Db 146 KKLNCYKITLECLPQNVGFYKKVGYT 171

RESULT 15
US-11-134-241-46
; Sequence 46, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
```


Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	199.5	23.9	165	6	US-10-449-902-34519	Sequence 34519, A
2	111	13.3	166	6	US-10-471-571A-3964	Sequence 3964, Ap
3	98	11.7	257	6	US-10-953-349-1260	Sequence 1260, Ap
4	96	11.5	254	6	US-10-449-902-29790	Sequence 29790, A
5	92	11.0	159	6	US-10-953-349-2743	Sequence 2743, Ap
6	92	11.0	223	6	US-10-953-349-2742	Sequence 2742, Ap
7	92	11.0	237	6	US-10-953-349-2741	Sequence 2741, Ap
8	88.5	10.6	149	6	US-10-471-571A-5116	Sequence 5116, Ap
9	83.5	10.0	139	6	US-10-370-959-55	Sequence 55, Appl
10	83.5	10.0	286	6	US-10-471-571A-2210	Sequence 2210, Ap
11	79.5	9.5	154	6	US-10-471-571A-260	Sequence 260, App
12	79.5	9.5	492	6	US-10-449-902-42958	Sequence 42958, A
13	79.5	9.5	492	6	US-10-449-902-52839	Sequence 52839, A
14	79	9.5	133	6	US-10-471-571A-4442	Sequence 4442, Ap
15	77	9.2	435	6	US-10-953-349-10209	Sequence 10209, A
16	77	9.2	603	6	US-10-953-349-10208	Sequence 10208, A
17	77	9.2	609	6	US-10-953-349-10207	Sequence 10207, A
18	75.5	9.0	158	6	US-10-471-571A-3244	Sequence 3244, Ap
19	75	9.0	317	6	US-10-449-902-31187	Sequence 31187, A
20	74.5	8.9	271	6	US-10-449-902-50947	Sequence 50947, A
21	74.5	8.9	466	6	US-10-471-571A-5066	Sequence 5066, Ap
22	74	8.9	711	6	US-10-471-571A-2458	Sequence 2458, Ap
23	73.5	8.8	183	6	US-10-471-571A-4252	Sequence 4252, Ap
24	73	8.7	251	6	US-10-953-349-26429	Sequence 26429, A
25	73	8.7	290	6	US-10-953-349-26428	Sequence 26428, A

Query Match 10.6%; Score 88.5; DB 6; Length 149;
Best Local Similarity 24.8%; Pred. No. 0.085;

	Query Match	10.0%	Score 83.5;	DB 6;	Length 139;
	Best Local Similarity	31.6%;	Pred.No.0.24;		
	Matches	18;	Conservative 15;	Mismatches 21;	Indels 3; Gaps 1;
Qy	95 GHIEDIAVNSKYOGGQGLKLLIDQLVTIGTGYDCYKIILDDCEKN---	VKFYEKCGF	148		
Db	59 GYLAMLAIVDSKYRRNGITGNLVKKAIIVAMVEGGDCEWLETTETNKSALKLYENLGF	115			


```
Query Match          9.5%; Score 79.5; DB 6; Length 492;
Best Local Similarity 27.5%; Pred. No. 2.8;
Matches 30; Conservative 15; Mismatches 55; Indels 9; Gaps 3;

QY 48 ATVMNDNEDKKIMQYNPMVIIDKR-TETVAATGNIIEERKIIHELGL-----CGH-IE 98
Db 23 APVMNDNEALTVGPRGPILLDDYHLIEKVAHFARERIPERVVHARGASAKGFFECTHDVT 82

QY 99 DIANSKYQGGGLKLLLDQLVITGFDYGCYKIIILDCDEKNVKFYKCG 147
Db 83 DITCADFLRSPCAQTPVIVRFSTVIHERGSPETIRDRGFAVKFYTRG 131

RESULT 14
US-10-471-571A-4442
; Sequence 4442, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4442
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(133)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-4442

Query Match          9.5%; Score 79; DB 6; Length 133;
Best Local Similarity 31.3%; Pred. No. 0.62;
Matches 26; Conservative 11; Mismatches 36; Indels 10; Gaps 2;

QY 72 TETVAATGNIIEERKIIHELGLCGHIEDIANSKYQGGGLKLLIDQL-----VTIGFD 125
Db 41 TVTLYDKDLRLGMGRVIGDGTVOIVDIANVLKSYQGGYGSLSINEHIMQYKGVAVEST 100

QY 126 YGCYKIIILDCDEKNVKFYKCGF 148
Db 101 Y----VSLIADYPADKLYTKGF 119

RESULT 15
US-10-953-349-10209
; Sequence 10209, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10209
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10209

Query Match          9.2%; Score 77; DB 6; Length 435;
Best Local Similarity 20.4%; Pred. No. 4.2;
Matches 30; Conservative 24; Mismatches 67; Indels 26; Gaps 4;
```

```
QY 10 RRMEEGDLEQVETLKVLTITTPESFSKLIKYWNEATVMNDNEDKKIMQYNPMVIID 69
Db 275 REAKVEDLAGIRQIIKPLEESGALVRRTDELLR-----ALDSFVVVE 317

QY 70 KRTETVAATGNI-IIEERKIIHELGLCGHIEDIANSKYQGGGLKLLIDQLVITGFDYGC 128
Db 318 REGHIIACAALFPFFEEK-----CGEVAATAVASDCRGQGGQKLLIDYIEKKASALGL 370

QY 129 YKIIILDCDEKNVKFYKCGFSNAGVEM 155
Db 371 EMLFL-LTTRTADWVRRRGFOECPIEM 396
```

Search completed: June 14, 2006, 15:34:54
Job time : 6.38281 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:28:00 ; Search time 14.9062 Seconds
(without alignments)
933.660 Million cell updates/sec

Title: US-10-612-779-30
Perfect score: 835
Sequence: 1 MSLPDGFYIRMEEGLEQV.....VXFYEKCGFSNAGVEMQIRK 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SID83/ptodata/2/iaa/5 COMB.pcp.*
- 2: /EMC_Celerra_SID83/ptodata/2/iaa/6 COMB.pcp.*
- 3: /EMC_Celerra_SID83/ptodata/2/iaa/7 COMB.pcp.*
- 4: /EMC_Celerra_SID83/ptodata/2/iaa/H COMB.pcp.*
- 5: /EMC_Celerra_SID83/ptodata/2/iaa/PCTUS COMB.pcp.*
- 6: /EMC_Celerra_SID83/ptodata/2/iaa/RE COMB.pcp.*
- 7: /EMC_Celerra_SID83/ptodata/2/iaa/baGkfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	100.0	159	2	US-09-538-092-266 Sequence 266, App
2	328	39.3	155	2	US-09-248-796A-14732 Sequence 14732, A
3	209	25.0	184	2	US-09-795-926-46 Sequence 46, Appl
4	209	25.0	184	2	US-10-364-774-46 Sequence 46, Appl
5	194	23.2	114	2	US-09-270-767-32856 Sequence 32856, A
6	134	23.2	114	2	US-09-270-767-48073 Sequence 48073, A
7	131.5	15.7	153	2	US-09-270-767-31849 Sequence 31849, A
8	131.5	15.7	153	2	US-09-270-767-47066 Sequence 47066, A
9	117.5	14.1	242	2	US-09-252-991A-28218 Sequence 28218, A
10	110	13.2	179	2	US-09-603-208A-20 Sequence 20, Appl
11	106	12.7	149	2	US-09-107-532A-5296 Sequence 5296, Ap
12	104	12.5	307	2	US-09-543-681A-6849 Sequence 6849, Ap
13	101	12.1	170	1	US-08-742-009-1 Sequence 1, Appli
14	101	12.1	170	2	US-09-172-110-1 Sequence 1, Appli
15	98.5	11.8	86	2	US-09-134-000C-6710 Sequence 6710, Ap
16	98	11.7	23	2	US-09-927-734C-5 Sequence 5, Appli
17	93	11.1	126	2	US-09-902-540-15328 Sequence 15328, A
18	91.5	11.0	205	2	US-09-107-532A-5462 Sequence 5462, Ap
19	91.5	11.0	227	2	US-09-328-352-4162 Sequence 4162, Ap
20	91.5	11.0	556	1	US-08-505-377-1 Sequence 1, Appli
21	91.5	11.0	556	2	US-08-798-269-1 Sequence 1, Appli
22	91.5	11.0	556	2	US-09-055-210-1 Sequence 1, Appli
23	91.5	11.0	556	2	US-09-298-924-8 Sequence 8, Appli
24	91	10.9	148	2	US-09-825-414-17 Sequence 17, Appl
25	90	10.8	145	2	US-09-107-532A-7085 Sequence 7085, Ap
26	89.5	10.7	204	2	US-09-489-039A-8937 Sequence 8937, Ap

27	89	10.7	140	2	US-09-489-039A-10789 Sequence 10789, A
28	89	10.7	470	2	US-09-248-796A-18961 Sequence 18961, A
29	89	10.7	1477	2	US-09-830-230A-414 Sequence 414, App
30	89	10.7	1494	2	US-09-830-230A-413 Sequence 413, App
31	88	10.5	172	2	US-09-543-681A-6412 Sequence 6412, Ap
32	88	10.5	180	2	US-09-107-532A-6074 Sequence 6074, Ap
33	86.5	10.4	240	2	US-09-489-039A-12910 Sequence 12910, A
34	86	10.3	150	2	US-09-543-681A-7022 Sequence 7022, Ap
35	85	10.2	235	2	US-09-543-681A-4771 Sequence 4771, Ap
36	84.5	10.1	161	2	US-09-603-208A-22 Sequence 22, Appl
37	84	10.1	148	2	US-09-489-039A-9177 Sequence 9177, Ap
38	83.5	10.0	362	2	US-10-104-047-3160 Sequence 3160, Ap
39	83	9.9	338	2	US-09-489-039A-12209 Sequence 12209, A
40	82.5	9.9	158	2	US-09-489-039A-9679 Sequence 9679, Ap
41	81.5	9.8	129	2	US-09-248-796A-17372 Sequence 17372, A
42	81.5	9.8	169	2	US-09-328-352-6116 Sequence 6116, Ap
43	81.5	9.8	237	2	US-09-252-991A-25273 Sequence 25273, A
44	81.5	9.8	442	2	US-09-489-039A-11693 Sequence 11693, A
45	80.5	9.6	168	2	US-08-990-791-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-538-092-266
; Sequence 266, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 266
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YFL017C
US-09-538-092-266

Query Match 100.0%; Score 835; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.2e-94;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLPDGFYIRMEEGLEQVTTKLTVTGTTTPEFSKLIKYWNEATVWNEDEKKIM 60
Db 1 MSLPDGFYIRMEEGLEQVTTKLTVTGTTTPEFSKLIKYWNEATVWNEDEKKIM 60
QY 61 QYNPMVIVDKRTVTAATGNIIERKIIHELGCHEIDIAVNSKYQGQGLGKLILDLV 120
Db 61 QYNPMVIVDKRTVTAATGNIIERKIIHELGCHEIDIAVNSKYQGQGLGKLILDLV 120
QY 121 TIGFDYGYCKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
Db 121 TIGFDYGYCKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159

RESULT 2
US-09-248-796A-14732
; Sequence 14732, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

Query Match 25.0%; Score 209; DB 2; Length 184;
Best Local Similarity 32.9%; Pred. No. 3.3e-17;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

```

; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32856
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32856

Query Match      23.2%; Score 194; DB 2; Length 114;
Best Local Similarity 40.2%; Pred. No. 1.1e-15;
Matches 35; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 62 YNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIANSKYQGGLGKLLIDQLVT 121
DB 17 YFTVIEDTRKNEIIGAASLVIERKFIHNRAVRGRLEDVVNDTYRGKQGLKLIIVTVSL 76

QY 122 IGFDYGCYKIILDCDKNVKFKYKCGF 148
DB 77 LAELGCGYKMSLDCKDKLIKFKYESLGY 103

RESULT 6
US-09-270-767-48073
; Sequence 48073, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 48073
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48073

Query Match      23.2%; Score 194; DB 2; Length 114;
Best Local Similarity 40.2%; Pred. No. 1.1e-15;
Matches 35; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 62 YNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIANSKYQGGLGKLLIDQLVT 121
DB 17 YFTVIEDTRKNEIIGAASLVIERKFIHNRAVRGRLEDVVNDTYRGKQGLKLIIVTVSL 76

QY 122 IGFDYGCYKIILDCDKNVKFKYKCGF 148
DB 77 LAELGCGYKMSLDCKDKLIKFKYESLGY 103

RESULT 7
US-09-270-767-31849
; Sequence 31849, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 31849
; LENGTH: 153
; TYPE: PRT

```

```

; ORGANISM: Drosophila melanogaster
US-09-270-767-31849

Query Match      15.7%; Score 131.5; DB 2; Length 153;
Best Local Similarity 29.6%; Pred. No. 7.5e-08;
Matches 34; Conservative 24; Mismatches 42; Indels 15; Gaps 4;

QY 9 IRRMEGDLQEQ-VTETLKVLTGTTGTTPESEKLIKVMNEATVMNNEDEKKIMQYNPMVI 67
DB 51 VRPLKDTDYDRGFLQLLSQLTHVGNVNRTOF--LTRF-----SQMKASGDYFVTVI 99

QY 68 VDKRTETVAATGNIIRKIIHELGLCGHIEDIANSKYQGGLGKLLIDQLVTI 122
DB 100 EDRKNEIIGAASLVIERKFIHNCAVRGRLEDVVNDTYRGKQGLKLIIV---VTV 151

RESULT 8
US-09-270-767-47066
; Sequence 47066, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47066
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47066

Query Match      15.7%; Score 131.5; DB 2; Length 153;
Best Local Similarity 29.6%; Pred. No. 7.5e-08;
Matches 34; Conservative 24; Mismatches 42; Indels 15; Gaps 4;

QY 9 IRRMEGDLQEQ-VTETLKVLTGTTGTTPESEKLIKVMNEATVMNNEDEKKIMQYNPMVI 67
DB 51 VRPLKDTDYDRGFLQLLSQLTHVGNVNRTOF--LTRF-----SQMKASGDYFVTVI 99

QY 68 VDKRTETVAATGNIIRKIIHELGLCGHIEDIANSKYQGGLGKLLIDQLVTI 122
DB 100 EDRKNEIIGAASLVIERKFIHNCAVRGRLEDVVNDTYRGKQGLKLIIV---VTV 151

RESULT 9
US-09-252-991A-28218
; Sequence 28218, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28218
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28218

Query Match      14.1%; Score 117.5; DB 2; Length 242;
Best Local Similarity 28.1%; Pred. No. 7.4e-06;
Matches 27; Conservative 22; Mismatches 44; Indels 3; Gaps 1;

```

```
Qy 65 MVIDKRTETVAATGNIIIEKIIHELGLCGHIEDIAVNSKYQGQGLKLLIDQLVTIGF 124
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 VTLVAERGERLLGCTCTLHLIEHLAHDFAARSAILDMVDRHARGQGVGRELIIGRAVERAR 203
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 125 DYGCYKILDCD---EKNVKPYEKGGSNAGVEMQI 157
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 SWGCYKLLALSSHQDRETAQREYAALGFTSHGVSLAL 239
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-603-208A-20
; Sequence 20, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: RESISTANCE AND TOLERANCE PROTEINS
; CURRENT APPLICATION NUMBER: US/09/603,208A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 20
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-20

Query Match 13.2%; Score 110; DB 2; Length 179;
Best Local Similarity 27.2%; Pred. No. 4e-05;
Matches 44; Conservative 23; Mismatches 69; Indels 26; Gaps 7;

Qy 7 FYIRMESGDLQVETELKVLTVTGTITPESFKLIKYNEATVWMDNEDKKIMQYNPMV 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 FTIRPIREGDPQVRDIYELGLETCHAT-----YETSGPTWDQFSQSKIND-TVMV 55
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 67 IVDKRTET-----VAATGNIIIEKIIHELGLCGHIED-IAVNSKYQGQGLKLLIDQLV 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 AVENNDPDPFILGWSAAP--ISSRQVFI-----GVVEDSIYIHPOGQGRGIGGALLDALI 108
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 TIGFDYGCVKI---ILDCDEKNVKPYEKGGSNAGVEMQIR 159
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 TYCSNGIWSHSWIFPENLGSAKLHESKGFVKVGTMQMAR 150
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 11
US-09-107-532A-5296
; Sequence 5296, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5296:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...149
; SEQUENCE DESCRIPTION: SEQ ID NO: 5296:
US-09-107-532A-5296

Query Match 12.7%; Score 106; DB 2; Length 149;
Best Local Similarity 36.1%; Pred. No. 9.4e-05;
Matches 26; Conservative 14; Mismatches 30; Indels 2; Gaps 1;

Qy 79 GNIIEEKIIHELGLCGHIEDIAVNSKYQGQGLKLLIDQLVTIGFDYGCVKI--ILDCD 136
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 GKIIIGVWITDFATFIQDILIHPRYQRQGIKGVLLNEALEKITSGPVQIELLTD 111
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 137 EKNVKPYEKGCGF 148
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 ETKKFEVSQVF 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-543-681A-6849
; Sequence 6849, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
```

```
US-08-742-009-1
```

		Query Match	12.1%;	Score 101;	DB 1;	Length 170;
		Best Local Similarity	23.1%;	Pred. No.	0.00046;	
Matches	37;	Conservative	26;	Mismatches	55;	Gaps 42; Indels 26;
QY	9	I RRM EGD LEO V TET KVL TTGCTTTPESF SKLIKYNEATVW DNDKDKIM OYNPMVIV	68			
Dd		:: :::: : :	:	:	:	: :
	6	I REAKEGGCDLRLRLELAEP-----EKLSDQVKISEELRALRDGF GDNPF--YHCLV--	56			
QY	69	DRKETVAATGNIIIEKKIHELGLCG-----HIEDIAVNKS YQGOG	110			
Dd	57	---AEILPAGKU-----LGPCVG VG I FYFYSTWKGRITYLEDIYVMPEYRGQG	104			
QY	111	L GKLLLDQLWTIGDFGYCYKI----ILDCEKNKVFEKGQ	147			
Dd	105	IGSKIIKKYAEEVALDKGCOSOFRLAVLDWMNORAMDLYKALG	144			

RESULT 14
US-09-172-110-1
; Sequence 1, Application US/09172110
; Patent No. 6017529
; GENERAL INFORMATION:
; APPLICANT: Hilman, Jennifer L.
; TITLE OF INVENTION: A NOVEL HUMAN SPERMIDINE/
; TITLE OF INVENTION: SPERMINE N1-ACETYLTRANSFERASE

```

ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/172,110
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/742,009
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0146 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE:
US-09-172-110-1

Query Match          12.1%; Score 101; DB 2; Length 170;
Best Local Similarity 23.1%; Pred.No. 0.00046;
Matches 37; Conservative 26; Mismatches 55; Indels 42; Gaps 6;

QY      9 IRRMERGDLQVETLKVLTITGTTPTSEFSKLIKYWNATVWVDNEDKKIMOYNPMIV 68
      || :||| : :: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB     6 IREAKEGDCGDILRLIRELAEF-----EKLSDQVKTISEEARADGFGNPF--YHCLV-- 56

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:21:37 ; Search time 10.7656 Seconds
(without alignments)
1421.048 Million cell updates/sec

Title: US-10-612-779-30
Perfect score: 835
Sequence: 1 MSLPDGFYIRRMEEGDLEQV.....VKFYEKCGFSNAGVEMQIRK 159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835	100.0	159	1 S56237	glucosamine-phosph
2	261.5	31.3	149	2 T51406	acetyltransferase-
3	232	27.8	111	2 T43426	glucosamine-phosph
4	214.5	25.7	165	2 T37319	probable acetyltra
5	192	23.0	347	2 T25192	hypothetical prote
6	125.5	15.0	157	2 AC1713	hypothetical prote
7	123	14.7	312	2 C71136	hypothetical prote
8	118.5	14.2	221	2 B83223	probable acetyltra
9	116.5	14.0	157	2 AE1342	hypothetical prote
10	113	13.5	154	2 G86699	hypothetical prote
11	113	13.5	166	2 E90074	conserved hypothet
12	112	13.4	140	2 C97300	probable acetyltra
13	108.5	13.0	146	2 E97249	probable acetyltra
14	106	12.7	185	2 AD1177	acetyltransferases
15	105	12.6	213	2 AH2054	N-terminal acetyl
16	105.5	12.6	195	2 AH1534	acetyltransferases
17	105	12.6	140	2 C97204	probable acetyltra
18	103.5	12.4	141	2 AF1484	weakly blastcidin
19	103	12.3	167	2 E97216	probable acetyltra
20	102.5	12.3	178	2 S75593	hypothetical prote
21	102	12.2	153	2 B83895	transcription regu
22	102	12.2	168	2 AD1534	B. subtilis regula
23	101	12.1	168	2 F83829	B. subtilis regula
24	100.5	12.0	144	2 A65018	hypothetical prote
25	100	12.0	178	2 A91042	hypothetical prote
26	100	12.0	178	2 D85886	hypothetical prote
27	100	12.0	284	2 G75045	hypothetical prote
28	100	12.0	284	2 G75045	probable acetyltra
29	99	11.9	177	2 T35580	probable acetyltra

ALIGNMENTS

RESULT 1

S56237

glucosamine-phosphate N-acetyltransferase (EC 2.3.1.4) [validated] - yeast (Saccharomyces cerevisiae)
N:Alternate names: phosphoglucosamine acetylase; phosphoglucosamine transacetylase; prot C:Species: Saccharomyces cerevisiae
C:Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text_change 09-Jul-2004
C:Accession: S56237; S48321; S62296

R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasano submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce
A:Reference number: S56186
A:Accession: S56237

A:Molecule type: DNA
A:Residues: 1-159 <MUR>

A:Cross-references: UNIPROT:P43577; UNIPARC:UPI000012B8BB; EMBL:D50617; NID:G936685; PID R:Churchev, C.
submitted to the EMBL Data Library, September 1994

A:Reference number: S48310
A:Accession: S48321

A:Molecule type: DNA
A:Residues: 1-111, 'ASS' <CHU>

A:Cross-references: UNIPARC:UPI0000168A6F; EMBL:Z46255; NID:G559925; PIDN:CAA86352.1; P R:Murakami, Y.
submitted to the EMBL Data Library, December 1994

A:Reference number: S62230
A:Accession: S62296

A:Molecule type: DNA
A:Residues: 1-159 <MUW>

A:Cross-references: UNIPARC:UPI000012B8BB; EMBL:D44596; NID:G1100783; PIDN:BAA08000.1; F C:Genetics:
A:Gene: SGD:GNA1; GNA1; MIPS:YFL017C

A:Cross-references: MIPS:YFL017C; SGD:S0001877
A:Map position: 6L

C:Function:
A:Description: EC 2.3.1.4 [validated, MUID:99085039]; glucosamine-phosphate N-acetyltran

A:Note: phosphoglucosamine acetyltransferase activity has been shown in vitro, by incuba -acetylglucosamine is produced from glucosamine 6-phosphate, indicating that 142-Phe and C:Superfamily: Saccharomyces glucosamine-phosphate N-acetyltransferase
C:Keywords: acyltransferase; coenzyme A

Query Match 100.0%; Score 835; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.3e-69;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLPDGFYIRRMEEGDLEQVTTTLKVLTVGTITPESFKLIKYNEATVWNDNEKIM 60
|||||

Db 1 MSLPDGFYIRRMEEGDLEQVTTTLKVLTVGTITPESFKLIKYNEATVWNDNEKIM 60
|||||

QY 61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLV 120
|||||

Db 61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLV 120
|||||

QY 121 TIGFDYGCYKIIILDCDEKNVFEKCGFSNAGVEMQIRK 159
 DB 121 TIGFDYGCYKIIILDCDEKNVFEKCGFSNAGVEMQIRK 159

RESULT 2
 T51406
 acetyltransferase-like protein - Arabidopsis thaliana
 N;Alternate names: protein F14P8_150
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C;Accession: T51406
 R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
 submitted to the Protein Sequence Database, August 2000
 A;Reference number: Z25394
 A;Accession: T51406
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-149 <SAT>
 A;Cross-references: UNIPROT:Q9LFU9; UNIPARC:UPI00000A553D; EMBL:AL391144
 A;Experimental source: cultivar Columbia; BAC clone F14P8
 C;Genetics:
 A;Map position: 5
 A;Note: F14P8_150
 C;Superfamily: Saccharomyces glucosamine-phosphate N-acetyltransferase

Query Match 31.3%; Score 261.5; DB 2; Length 149;
 Best Local Similarity 35.4%; Pred. No. 9.3e-17;
 Matches 56; Conservative 32; Mismatches 51; Indels 19; Gaps 3;

QY 3 LPDGFYIRRMEDGLEQ-VTEFLKVLTVGTITPESFSLIKIYWEATVWVNDNEDKKIMQ 61
 DB 1 MAETFKIRKLEISDKRGFIQLGLTVGTSVTDDEP-----DRRFEIRS 46

QY 62 YNP----MVIYDKRTETVAATGNIIIRKIIHEGLGCHIEDIIVNSKYQGGLGLLIDLV 117
 DB 47 YGDDHVICVEETSGKIATGSVMIEKFLRNGKAGHIEDVVVDSFRFRGKQLGKKVVE 106

QY 118 QLVITGFDYGCYKIIILDCDEKNVFEKCGFSNAGVEM 155
 DB 107 FLMDHCKSMGCKYKVLDCSVENKVFYKCGMSNKSIO 144

RESULT 3
 T43426
 glucosamine-phosphate N-acetyltransferase (EC 2.3.1.4) [similarity] - fission yeast (Sch
 C;Species: Schizosaccharomyces pombe
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C;Accession: T43426; T37783
 R;Mio, T.; Yamada-Okabe, T.; Arisawa, M.; Yamada-Okabe, H.
 J. Biol. Chem. 274, 424-9, 1999
 A;Title: Saccharomyces cerevisiae GNAL, an essential gene encoding a novel acetyltransfe
 A;Reference number: Z14069; PMID:9867860; PMID:9867860
 A;Accession: T43426
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-111 <MIO>
 A;Cross-references: UNIPROT:O13738; UNIPARC:UPI000013AA1D; EMBL:AB017629; PIDN:BAA36498.
 R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1995
 A;Reference number: Z21746
 A;Accession: T37783
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-111 <OLI>
 A;Cross-references: UNIPARC:UPI000013AA1D; EMBL:Z98529; PIDN:CAB11032.1; GSPDB:GN000066;
 A;Experimental source: strain 972h; cosmid c16E8
 C;Genetics:
 A;Gene: GNAL; SPAC16E8.03
 A;Map position: 1
 A;Introns: 22/3
 C;Keywords: acyltransferase; coenzyme A

Query Match 27.8%; Score 232; DB 2; Length 111;
 Best Local Similarity 42.1%; Pred. No. 3.3e-14;
 Matches 40; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 62 YNPWIVDKRTETVAATGNIIIRKIIHEGLGCHIEDIIVNSKYQGGLGLLIDLV 121
 DB 9 YYIIVVEDLSEHHVIGTATFLERKFLRGKGLGCHIEEIVHPDHRKAIGKLVLTLLK 68

QY 122 IGFDYGCYKIIILDCDEKNVFEKCGFSNAGVEMQ 156
 DB 69 LAFSLNSYKVIILDCSDSNVGFYKCGLSRAGIEMK 103

RESULT 4
 T37319
 probable acetyltransferase (EC 2.3.1.-) - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T37319; T18647
 R;Okabe, H.
 submitted to the EMBL Data Library, September 1998
 A;Reference number: Z21684
 A;Accession: T37319
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-165 <OKA>
 A;Cross-references: UNIPROT:Q17427; UNIPARC:UPI000012B8B9; EMBL:AB017628; PIDN:BAA36497.
 R;McMurray, A.
 submitted to the EMBL Data Library, April 1996
 A;Reference number: Z19001
 A;Accession: T18647
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-165 <WIL>
 A;Cross-references: UNIPARC:UPI000012B8B9; EMBL:Z71178; PIDN:CAA94884.1; GSPDB:GN000023;
 A;Experimental source: clone B0024
 C;Genetics:
 A;Gene: GNAL; B0024.12
 A;Map position: 5
 C;Superfamily: Saccharomyces glucosamine-phosphate N-acetyltransferase
 C;Keywords: acyltransferase

Query Match 25.7%; Score 214.5; DB 2; Length 165;
 Best Local Similarity 30.7%; Pred. No. 2.1e-12;
 Matches 46; Conservative 36; Mismatches 57; Indels 11; Gaps 2;

QY 2 SLDPGFYIRRMEDGLEQ-VTEFLKVLTVGTITPESFSLIKIYWEATVWVNDNEDKKIM 60
 DB 17 NLPDNFKVRPLAKDDFSKGVVDLLSQLTSVGNLDQEAPEKRFEMRTS-----VP 66

QY 61 QYNPWIVDKRTETVAATGNIIIRKIIHEGLGCHIEDIIVNSKYQGGLGLLIDLV 120
 DB 67 NYHIVWIEDSNSQKVASASLWEMKFIHAGSGRVEDVVVDTEMRQKLGAVLLTLV 126

QY 121 TIGFDYGCYKIIILDCDEKNVFEKCGFSN 150
 DB 127 SLGSLGVYKISLECVPELLPFYQFGFD 156

RESULT 5
 T25192
 hypothetical protein T23G11.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T25192
 R;Gardner, A.
 submitted to the EMBL Data Library, October 1996
 A;Reference number: Z19993
 A;Accession: T25192
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-347 <WIL>
 A;Cross-references: UNIPROT:O45811; UNIPARC:UPI0000077DB9; EMBL:Z81130; PIDN:CAB03416.1;

A; Gene: CAC2840
C; Superfamily: Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimi

```

QY 00 KLNH0000GHLELVAHNSNAKQVSTGTMUMLQVYVGF02SCNIAADDQENW-----149
Db 63 -IIDE-----GHITNIAVHPFRGNGIGNKILEELKL-----CEKNIPSWTL 105
QY 141 -----KFYEKCGFSNAGV 153
Db 106 EVRISNTIAQNLYKKFGFKEAGV 128

```

Search completed: June 14, 2006, 15:28:35
Job time : 12.7656 secs

Qy 133 LDCDEKN---VKFYEKCGFSNAG 152
 |||:|:|:|:
 Db 149 LNCDCKGNPHAKRLYERLGFHVTG 171

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:15:09 ; Search time 79.707 Seconds
(without alignments)
1845.226 Million cell updates/sec

Title: US-10-612-779-30

Perfect score: 835

Sequence: 1 MSLPDGFYIRMEEGDLEQV.....VKFYKCGFSGNAGVEMQIRK 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	835	100.0	159	1	GNAL_YEAST	P43577 saccharomyc
2	476.5	57.1	160	2	Q6FW86 CANGA	Q6fw86 candida gla
3	434.5	52.0	171	2	Q752R3 ASHGO	Q752r3 ashya goss
4	430.5	51.6	188	2	Q6CSR3 KLULA	Q6csr3 kluyveromyc
5	347	41.6	150	2	Q6BX84 DEBHA	Q6bx84 debaryomyc
6	336	40.2	171	2	Q6C8F2 YARLI	Q6c8f2 yarrowia li
7	333	39.9	149	2	Q5AHF9 CANAL	Q5ahf9 candida alb
8	328	39.3	149	1	GNAL_CANAL	Q93806 candida alb
9	314	37.6	170	2	Q2UGQ9 ASPOR	Q2ugq9 aspergillus
10	290	34.7	157	2	Q54WR8 DICDI	Q54wr8 dictyosteli
11	279.5	33.5	177	2	Q7SHA4 NEUCR	Q7sha4 neurospora
12	276	33.1	176	2	Q411G4 GIBZE	Q4ilg4 gibberella
13	275	32.9	190	2	Q4WCU5 ASPFU	Q4wcus aspergillus
14	266	31.9	165	2	Q5SR34_CRYNE	Q5sr34 cryptococcu
15	266	31.9	165	2	Q5KF35_CRYNE	Q5kf35 cryptococcu
16	264	31.6	205	2	Q7QTX4 GIALA	Q7qtx4 giardia lam
17	264	31.6	205	2	Q86F18 GIALA	Q86f18 giardia lam
18	261.5	31.3	149	2	Q9LFU9 ARATH	Q9lfu9 arabidopsis
19	249	29.8	177	2	Q4IE15_GIBZE	Q4ie15 gibberella
20	246.5	29.5	181	2	Q2KFU6_MAGGR	Q2kfu6 magnaporthe
21	236	28.3	148	2	Q7R0Y7 GIALA	Q7r0y7 giardia lam
22	235	28.1	159	2	Q5ONZ5_ENTHI	Q5onz5 entamoeba h
23	234.5	28.1	158	2	Q51AE1_ENTHI	Q51ae1 entamoeba h
24	232	27.8	111	1	GNAL_SCHPO	Q13738 schizosacch
25	226	27.1	234	2	Q5ASM4_EMENI	Q5asm4 aspergillus
26	222.5	26.6	333	2	Q4PLQ6_USTWA	Q4plq6 ustilago ma
27	216	25.9	203	2	Q8SRG4_ENCCU	Q8sr94 encephalito
28	215	25.7	182	2	Q86QP1_BRABE	Q86qp1 brachydanio
29	214.5	25.7	165	1	GNAL_CAEEL	Q17427 caenorhabdi
30	209.5	25.1	1100	2	Q5KF71_CRYNE	Q5kf71 cryptococcu
31	209.5	25.1	1127	2	Q55QZ7_CRYNE	Q55qz7 cryptococcu

32	208.5	25.0	179	2	Q4SAJ2_TETNG	Q4saj2 tetraodon n
33	208	24.9	159	2	Q5ORB0_ENTHI	Q5orb0 entamoeba h
34	208	24.9	184	1	GNAL_HUMAN	Q96ek6 homo sapien
35	208	24.9	184	1	GNAL_PONPY	Q5ral9 pongo pygma
36	207	24.8	215	2	Q8IMK5_DROME	Q8imk5 drosophila
37	207	24.8	219	1	GNAL_DROME	Q9vai0 drosophila
38	206	24.7	148	2	Q4DGL9_TRYCR	Q4dgl9 trypanosoma
39	206	24.7	148	2	Q4DX16_TRYCR	Q4dx16 trypanosoma
40	205.5	24.6	148	1	GNAL_MIMIV	Q5upz9 mimivirus
41	204	24.4	184	2	Q4VBJ4_BRARE	Q4vbj4 brachydanio
42	201.5	24.1	198	2	Q7Q212_ANOGA	Q7q212 anopheles g
43	201	24.1	184	1	GNAL_MOUSE	Q9jqk8 mus musculu
44	201	24.1	349	2	Q6OW31_CARBR	Q6ow31 caenorhabdi
45	200	24.0	212	2	Q5Q126_AEDAE	Q5q126 aedes aegyp

ALIGNMENTS

RESULT 1

GNAL_YEAST	ID	GNAL_YEAST	STANDARD;	PRT;	159 AA.
AC	P43577;				
DT	01-NOV-1995,	integrated into UniProtKB/Swiss-Prot.			
DT	01-NOV-1995,	sequence version 1.			
DT	07-MAR-2006,	entry version 52.			
DE	Glucosamine 6-phosphate N-acetyltransferase (EC 2.3.1.4)				
DE	(Phosphoglucosamine transacetylase) (Phosphoglucosamine acetylase).				
GN	Name=GNAL; Synonyms=PAT1; OrderedLocuNames=YFL017C;				
OS	Saccharomyces cerevisiae (Baker's Yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
OX	NCBI_TaxID=4933;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA], CHARACTERIZATION, AND MUTAGENESIS.				
RX	MEDLINE=99085039; PubMed=9867860; DOI=10.1074/jbc.274.1.424;				
RA	Mio T., Yamada-Okabe T., Arisawa M., Yamada-Okabe H.;				
RT	"Saccharomyces cerevisiae GNAL an essential gene encoding a novel acetyltransferase involved in UDP-N-acetylglucosamine synthesis.";				
RL	J. Biol. Chem. 274:424-429(1999).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=S288c / AB972;				
RX	MEDLINE=95400292; PubMed=7670463;				
RA	Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,				
RA	Sasanuma S., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,				
RA	Yamazaki M., Tashiro H., Eki T.;				
RT	"Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae.";				
RL	Nat. Genet. 10:261-268(1995).				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].				
RC	STRAIN=S288c / AB972;				
RA	Bartell B.G., Churche C., Rajandream M.A.;				
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].				
RC	STRAIN=S288c;				
RA	Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,				
RA	Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,				
RA	Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,				
RA	Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,				
RA	LaBaer J.;				
RT	"Creation of the YFLEX clone resource: cloning of Saccharomyces cerevisiae ORFs in the Gateway recombinational cloning system.";				
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	LEVEL OF PROTEIN EXPRESSION				
RX	MEDLINE=22923965; PubMed=14562106; DOI=10.1038/nature02046;				
RA	Ghaemmaghami S., Huh W.-K., Bower K., Howson R.W., Belle A.,				
RA	Dephour N., O'Shea E.K., Weissman J.S.;				
RT	"Global analysis of protein expression in yeast.";				
RL	Nature 425:737-741(2003).				


```

Db      121 TEIGFDAGCYKVILDCDEKNVAFYEKCYKAGVEMQCR 159

RESULT 3
Q752R3 ASHGO PRELIMINARY; PRT; 171 AA.
AC Q752R3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE AFR510Wp.
GN OrderedLocusNames=AFR510W;
OS Ashbya gossypii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehmann R., Luedi P., Choi S., Wing R.A., Flavler A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF016819; AAS53881.1; -; Genomic_DNA.
DR AGD; AFR510W; -;
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR InterPro; IPR000182; GCM5acetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome.
SQ SEQUENCE 171 AA; 18855 MW; 69F2E5310C2A186D CRC64;

Query Match 52.0%; Score 434.5; DB 2; Length 171;
Best Local Similarity 56.8%; Pred. No. 2.7e-31;
Matches 92; Conservative 16; Mismatches 47; Indels 7; Gaps 2;

QY 3 LPDGFVIRMEEGDLEQVETLKVLTGVTITPESFKLIKYWNEATVWVNDNEDKK----58
Db 4 LPEGHVIRAEADYAGVETLKVLTGVDGTEREFAERIAIYKTVKVPVARGKRPVGM 63
QY 59 --IMQNPMPVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYQGGLGLKLI 116
Db 64 GEILAYNPVITDE-AGRVATGNIIRKIIHELGLCGHIEDIIVASDQQKRLGLMLI 122
QY 117 DQLVITGFYGYKIIILDCDEKNVAFYEKCYKAGVEMQCR 158
Db 123 NTLTEIRNAGCYKIIILDCDQPNADFYKCKGFSQAGLEMQHR 164

RESULT 4
Q6CRN3 KLULA PRELIMINARY; PRT; 188 AA.
AC Q6CRN3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome D of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN OrderedLocusNames=KLIA0D07700g;
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;

RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Blanchin S., Beckerich J.-M., Anthouard V., Babour A., Barbe V.,
RA Boisarame A., Boyer J., Cattolico L., Confaniolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR382124; CAH00502.1; -; Genomic_DNA.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR InterPro; IPR000182; GCM5acetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome.
SQ SEQUENCE 188 AA; 21116 MW; ED656D6A8C849071 CRC64;

Query Match 51.6%; Score 430.5; DB 2; Length 188;
Best Local Similarity 53.8%; Pred. No. 6.9e-31;
Matches 85; Conservative 21; Mismatches 45; Indels 7; Gaps 2;

QY 1 MSLPDGFVIRMEEGDLEQVETLKVLTGVTITPESFKLIKYWNEATVWVNDNEDKIM 60
Db 35 MSLPEGYTIIRTKKDDFGVTSVLKVTWGVDSQDFHSLDHDVSTIGD----IP 88
QY 61 QYNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYQGGLGLKLI 120
Db 89 MYSNTIVDTNGE-VVATGNVLVEKVIHECGLVGHIEDIIVAKDQKGLGLLIQYLY 147
QY 121 TGFYDYGCVKIIILDCDEKNVAFYEKCYKAGVEMQCR 158
Db 148 KLANEYGCYKIIILDCDESNVGFYKCGLKKAGVEMQIR 185

RESULT 5
Q6BX84 DEBHA PRELIMINARY; PRT; 150 AA.
AC Q6BX84;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Similar to sp|O93806|Candida albicans Glucosamine 6-phosphate N-
DE acetyltransferase.
GN OrderedLocusNames=DEHA0B05005g;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36239 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Blanchin S., Beckerich J.-M., Anthouard V., Babour A., Barbe V.,
RA Boisarame A., Boyer J., Cattolico L., Confaniolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
```

RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Boucher C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.,
RT "Genome evolution in yeasts";
RL Nature 430:35-44(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC ENBL; CR382134; CAG85180.1; -: Genomic DNA.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000182; GCS5acetyl_trans.
DR Pfam; PF00583; Acetyltransf 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 150 AA; 16926 MW; 5BF8B8B7DE0C98B7C8 CRC64;

Query Match 41.6%; Score 347; DB 2; Length 150;
Best Local Similarity 48.1%; Pred. No. 2.1e-23;
Matches 75; Conservative 17; Mismatches 52; Indels 12; Gaps 3

QY 1 MSLPDGVFIRRMERGDL-EQVTEITKVLTTVGTITPESFSLIKYVNEATVWNNEDKKI 59
DB 1 MQLPEGYSFRLQNDQVSNVYLETIKVLTTVGEISPELFAVFNWQSLP ----- 50
QY 60 MQYNPMVIVDKRTETVAATGNIIIERKIIHBLGCGHIEDIAVNSKYQGQGLGKLIDQL 119
DB 51 EIIYQPHVITN-NDGTVATGMLFIERKVIHECGSVGHIEDIAVAKSEQGKLGFSWISGL 109
QY 120 VTIGFDYGCYKIILDCDEKNVKFYEKCGFSNAGVEM 155
DB 110 TEVAKNGCYKIILDCSPHNKVFYKCGYKNDGVEM 145

RESULT 6
ID Q6C8F2_YARLI PRELIMINARY; PRT; 171 AA.
AC Q6C8F2;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 15-AUG-2004, sequence version 1.
DT 17-FEB-2006, entry version 12.
DE Similar to sp|O93806 Candida albicans Glucosamine 6-phosphate N-
DE acetyltransferase.
GN OrderedLocNames=YALI0D20152g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
NCBI_TaxID=4952;
RX [1]
RN NCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIB 122 / E 150;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Tallia E.,
RA Goffard N., Franguel L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boixame A., Boyer J., Cattolico L., Confalonieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Boucher C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts";
RL Nature 430:35-44(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC ENBL; CR382130; CAG81252.1; -: Genomic DNA.

[illegible]

```

CC -----
DR EMBL; AAHF01000012; EAL85793.1; -; Genomic DNA.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR KW Complete proteome; Transferase.
SQ SEQUENCE 190 AA; 21102 MW; F72COFFCDD0E40A3 CRC64;

Query Match 32.9%; Score 275; DB 2; Length 190;
Best Local Similarity 41.3%; Pred. No. 9.9e-17;
Matches 64; Conservative 19; Mismatches 58; Indels 14; Gaps 5;

Qy 3 LPDGFYIRRMEEGDLEQ-VVETLKLVLTTVGTITPESFSKLIKYNEATVW-NDNEDKKIM 60
Db 44 LPADYTRPCLRSDDYKRGYLDVLRLVLTGVDINEQ-----WNSRYEWIRARSD--- 92

Qy 61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCHGHIEDIAVNSKYQGQGLKLLDQLV 120
Db 93 EYTLVAVCDCEGR-IVGTGSLVVERKFIHSLGMVGHIEDIAVKGQOGKGLGRILIQALD 151

Qy 121 TIGFDGVCYKIIILDCDEKNKVFYKCGFSNAGVEM 155
Db 152 YVAEKVGCYKTIILDCSEANEGFYIKCGFKRAGLEM 186

RESULT 14
Q5SR34_CRYNE PRELIMINARY; PRT; 165 AA.
ID Q5SR34_CRYNE
DC Q5SR34;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DR Hypothetical protein.
GN ORFNames=CNBF1560;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
ON NCBI_TaxID=283643;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AASY01000030; EAL20345.1; -; Genomic DNA.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR KW Hypothetical protein.
SQ SEQUENCE 165 AA; 18054 MW; 415C613042BD20B2 CRC64;

Query Match 31.9%; Score 266; DB 2; Length 165;
Best Local Similarity 43.2%; Pred. No. 5.6e-16;
Matches 63; Conservative 22; Mismatches 49; Indels 12; Gaps 5;

Qy 4 PDGFYIRRMEEGD-LEQVETLKLVLTTVGTITPESFSKLIKYNEATVWNDNEDKKIMQY 62
Db 26 PD-LYRPLSSSTDVLRGHIELLSVLTA---PPQSVS-----TYETIFQMKASAGI-Y 74

Qy 63 NPMMVIVDKRTETVAATGNIIIERKIIHELGLCHGHIEDIAVNSKYQGQGLKLLDQLV 122
Db 75 FTVVVHRLSNQVAVCGSVIIERKFVRNAGLVGHIEDIAVSQSNQGRKLGKMIINTLVDI 134

Qy 123 GFYDVCYKIIILDCDEKNKVFYKCGF 148
Db 135 GLARGCYKIIILDCSEKNIPFEKCY 160

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:30:10 ; Search time 20.6172 Seconds
(without alignments)
666.288 Million cell updates/sec

Title: US-10-612-779-6
Perfect score: 3089
Sequence: 1 MCGIVCAIAQRDVAEILLEG.....LIKGTVDQPRNLAKSVTVYE 609

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 96747 seqs, 22556637 residues

Total number of hits satisfying chosen parameters: 96747

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /EMC_Celerra_SID33/ptodata/2/pubpaa/US09 NEW PUB.pap.*
2: /EMC_Celerra_SID33/ptodata/2/pubpaa/US06 NEW PUB.pap.*
3: /EMC_Celerra_SID33/ptodata/2/pubpaa/US07 NEW PUB.pap.*
4: /EMC_Celerra_SID33/ptodata/2/pubpaa/US08 NEW PUB.pap.*
5: /EMC_Celerra_SID33/ptodata/2/pubpaa/FCT NEW PUB.pap.*
6: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10 NEW PUB.pap.*
7: /EMC_Celerra_SID33/ptodata/2/pubpaa/US11 NEW PUB.pap.*
8: /EMC_Celerra_SID33/ptodata/2/pubpaa/US60_NEW_PUB.pap.*

* Pred. No. is the number of results predicted by chance to have a
* score greater than or equal to the score of the result being printed,
* and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3089	100.0	609	7 US-11-245-473-22	Sequence 22, Appl
2	3072	99.4	609	7 US-11-245-473-28	Sequence 28, Appl
3	3072	99.4	609	7 US-11-245-473-31	Sequence 31, Appl
4	3068	99.3	609	7 US-11-245-473-16	Sequence 16, Appl
5	3061	99.1	609	7 US-11-245-473-25	Sequence 25, Appl
6	3053	98.8	609	7 US-11-245-473-19	Sequence 19, Appl
7	1145	37.1	601	6 US-10-471-571A-2268	Sequence 2268, Ap
8	1104	35.7	681	6 US-10-480-021-6	Sequence 6, Appl
9	1104	35.7	681	6 US-10-480-021-7	Sequence 7, Appl
10	1098.5	35.6	682	6 US-10-480-021-8	Sequence 8, Appl
11	1078.5	34.9	680	6 US-10-953-349-34174	Sequence 34174, A
12	1071	34.7	697	6 US-10-449-902-38427	Sequence 38427, A
13	801	25.9	491	6 US-10-953-349-34175	Sequence 34175, A
14	800	25.9	488	6 US-10-953-349-34176	Sequence 34176, A
15	582.5	18.9	464	6 US-10-488-015-17	Sequence 17, Appl
16	551.5	17.9	286	6 US-10-449-902-36965	Sequence 36965, A
17	252.5	8.2	541	6 US-10-449-902-54059	Sequence 54059, A
18	241	7.8	501	6 US-10-471-571A-5078	Sequence 5078, Ap
19	190.5	6.2	600	6 US-10-449-902-38704	Sequence 38704, A
20	184.5	6.0	78	6 US-10-488-015-16	Sequence 15, Appl
21	156.5	5.1	591	6 US-10-449-902-53151	Sequence 53151, A
22	152.5	4.9	584	6 US-10-953-349-4929	Sequence 4929, Ap
23	143	4.6	54	6 US-10-488-015-18	Sequence 18, Appl
24	140	4.5	582	6 US-10-953-349-22931	Sequence 22931, A
25	134	4.3	73	6 US-10-488-015-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-11-245-473-22
; Sequence 22, Application US/11245473
; Publication No. US2006094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; CURRENT FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli

US-11-245-473-22

Query Match 100.0%; Score 3089; DB 7; Length 609;

Best Local Similarity 100.0%; Pred. No. 8.7e-221;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MCGIVCAIAQRDVAEILLEGRLRYGVDSAGLAVVDTEGHMTRRLRLKQVQMLAAAE	60
Db	1	MCGIVCAIAQRDVAEILLEGRLRYGVDSAGLAVVDTEGHMTRRLRLKQVQMLAAAE	60
Qy	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIENHPELREELKARGYTFV	120
Db	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIENHPELREELKARGYTFV	120
Qy	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGPLV	180
Db	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGPLV	180
Qy	181	IGLGMENFTASDQALLPVTRRFIIEGDIAREITRSVNIIDKTAEVKRODIESNLQ	240
Db	181	IGLGMENFTASDQALLPVTRRFIIEGDIAREITRSVNIIDKTAEVKRODIESNLQ	240
Qy	241	YDAGDKGIYCHVMQKEIYEQPNAIKNTLTGRISHGVQDLSLGNADLKSKEVHIOLA	300
Db	241	YDAGDKGIYCHVMQKEIYEQPNAIKNTLTGRISHGVQDLSLGNADLKSKEVHIOLA	300

Sequence 5406, Ap
Sequence 326, App
Sequence 50327, A
Sequence 37184, A
Sequence 24311, A
Sequence 24310, A
Sequence 43, Appl
Sequence 37185, A
Sequence 54230, A
Sequence 54230, A
Sequence 36548, A
Sequence 36548, A
Sequence 41204, A
Sequence 27, Appl
Sequence 32217, A
Sequence 32216, A
Sequence 8931, Ap
Sequence 52284, A
Sequence 32551, A
Sequence 2, Appli

554 6 US-10-471-571A-5406
586 7 US-11-289-102-326
797 6 US-10-449-902-50327
604 6 US-10-953-349-37184
461 6 US-10-953-349-24311
492 6 US-10-953-349-24310
866 7 US-11-121-154-43
579 6 US-10-953-349-37185
601 6 US-10-449-902-54230
790 6 US-10-449-902-50553
455 6 US-10-449-902-36548
562 6 US-10-953-349-37186
1045 6 US-10-449-902-41204
60 6 US-10-488-015-27
527 6 US-10-953-349-32217
542 6 US-10-953-349-32216
437 6 US-10-953-349-8931
577 6 US-10-449-902-52284
831 6 US-10-953-349-32551
4074 6 US-10-501-834-2

26 116 3.8
27 116 3.8
28 107 3.5
29 106.5 3.4
30 105 3.4
31 105 3.4
32 105 3.4
33 102.5 3.3
34 102.5 3.3
35 102.5 3.3
36 101 3.3
37 101 3.3
38 101 3.3
39 100.5 3.3
40 100.5 3.3
41 100.5 3.3
42 100 3.2
43 99 3.2
44 98 3.2
45 98 3.2

Db 241 YDAGDKGIYCHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CQTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Db 301 CQTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLSRGQOYPIA 480
Qy 481 LEGALKKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLEKSNIEVRAR 540
Db 481 LEGALKKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLEKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

US-11-245-473-28
; Sequence 28, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-11-245-473-28

Query Match 99.4%; Score 3072; DB 7; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.6e-219;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MCGIVGAIQAQDVABEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60
Db 1 MCGIVGAIQAQDVABEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPVL 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPVL 180
Qy 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240

Qy 241 YDAGDKGIYCHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CQTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Db 301 CQTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLSRGQOYPIA 480
Qy 481 LEGALKKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLEKSNIEVRAR 540
Db 481 LEGALKKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLEKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3

US-11-245-473-31
; Sequence 31, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; CURRENT FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-11-245-473-31

Query Match 99.4%; Score 3072; DB 7; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.6e-219;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MCGIVGAIQAQDVABEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60
Db 1 MCGIVGAIQAQDVABEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPVL 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPVL 180
Qy 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240

Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRWYFESLAGIPCDEVIASEPRYKSAVRNLSIMITLSOSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWYFESLAGIPCDEVIASEPRYKSAVRNLSIMITLSOSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEDEFSKHALFLSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEDEFSKHALFLSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 6

US-11-245-473-19
; Sequence 19, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-19

Query Match 98.8%; Score 3053; DB 7; Length 609;

Best Local Similarity 99.0%; Pred. No. 4e-218; Mismatches 6; Indels 0; Gaps 0;
Matches 603; Conservative 0;

QY 1 MCGIVGAIQAORDVAREILLEGRLRLEYRGYDVSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 60
Db 1 MCGIVGAIQAORDVAREILLEGRLRLEYRGYDVSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 60
QY 61 EHPLHGCTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHNEPREELKARGYTFV 120
Db 61 EHPLHGCTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHNEPREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRWYFESLAGIPCDEVIASEPRYKSAVRNLSIMITLSOSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWYFESLAGIPCDEVIASEPRYKSAVRNLSIMITLSOSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEDEFSKHALFLSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEDEFSKHALFLSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTVDVQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 7

US-10-471-571A-2268
; Sequence 2268, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2268
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2268

Query Match 37.1%; Score 1145; DB 6; Length 601;

Best Local Similarity 40.6%; Pred. No. 4.8e-77;
Matches 251; Conservative 126; Mismatches 215; Indels 26; Gaps 10;

QY 1 MCGIVGAIQAORDVAREILLEGRLRLEYRGYDVSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 60
Db 1 MCGIVGAIQAORDVAREILLEGRLRLEYRGYDVSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 59
QY 61 EHPLHGCTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHNEPREELKARGYTFV 118
Db 60 SSDPDPVGIHTRWATHGVPNHSHPHQSSNGRFTLVHNGVIEYELKGEY-LQGVS 118
QY 119 FVSTDDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLLAARSGSP 178
Db 119 FVSTDDTEVIAHLVNWELKQGGTLREAVLRAI POLRGAYCTVIMDSRHPDPTLLAARSGSP 178

Db	60	RYAGAPLVPRQEGKIENLVRSVYSEVDEKVDNLDAFVSH--AGIAHTRWATEGV	117
Qy	84	VNAHPHVS---EHI VVHNGIIEHPELREELKARGYTFVSETDTEVIAHLVNWEL----	136
Db	118	RNSHPQSSGAGDEFVLVHNGIITNYEVLKETLTRHGFTFESTDTEVTPKLAKFPVDKSH	177
Qy	137	-KQGG-TLREAVILRAIPOLRGAYGTVMDSRHPDTLLAARSGSPVLVIGL-----	184
Db	178	DEQGDVTFSQVVMVMEVRLGEGALYIFKSPHYNELIAKRGSQILGVNELSGQONGKS	237
Qy	185	-----MGENFTASDQALLPVTTRPFI FLEEGDIAEITRRSVNI--FD---KTG	227
Db	238	FHDVKLTITNGKPKELPFSSDLCAIVEHTKNVLALEDNEIVHI KQGSVILKFDPHKBP	297
Qy	228	AEVKRQDIESLOYDAGDKGKYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGNAD	287
Db	298	ASVQRALSVLSEMEVEQIKKGSYDHFMQKEIIEHQPHSLKTTWGRKLKDGQWLVGGKEYVL-	356
Qy	288	ELLSKVEHIQTLACCTSYNSGMVSRWYFESLAGIPCDVEIASEFERYKRSAVRNSIMITL	347
Db	357	KTIRRCRRVVGCGTSYNAALARPFEELTGTIPVTMEVASDLLDRQGPYREDTAVFV	416
Qy	348	SQSGETADTLAGLRSLKELGYLGSIAICNPQSSLVRESDLAMTNAGTEIGVASTKAPT	407
Db	417	SQSGETADTLAALDYALENGAL-CVGI TNTVGSTLSRKTCHGWHINAGCEIGVASTKAYT	475
Qy	408	TOLTVLMLVAKLSRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEDFSXHH	467
Db	476	SQIVAMAMMALAIGSDQISTQARRDSIISGLNLSNVSEVLKLDAGMKELASSLIDSES	535
Qy	468	ALFTLSRGDQYPIALEGALKLKEISYIHAEVYAAGELKHGKPLALIDADMPVIVVAPNNELL	527
Db	536	LLVFCRGYNATALEGALKVREVALMHSEGLMAGEMKHGKPLAVDENLPIIVIAITRDACF	595
Qy	528	EKLKSNTIEEVRARGQLVVFADQ-DAGFV---SSDNNKHIIEMPHVEVIAPIFYTVPLQLL	584
Db	596	SKQSVTQQLLSRRGRLLVMSRGDASAVCPSGGSGCRVIEVPQVADCLQPVINIPLQLL	655
Qy	585	AYHVALIKGTVDQDPNNLAKSVTVE 609	
Db	656	AYHLTVLRGFDVDQDPNNLAKSVTTQ 680	

```

RESULT 12
US-10-449-902-38427
; Sequence 38427, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203369
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38427
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-38427

```

[illegible]

```

RESULT 13
US-10-953-349-34175
; Sequence 34175, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34175
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34175

```



```

Qy 324 DVEIASEFRYKSAVRNSLMTLSOSGETADTLAGIRLSKELGYLGSLAICNVPGSSLV 383
Db 306 SVELASDFLDQAPVFRDDTCVFSQSGETADSLMALRYCLERGA-TVGINNVGSSIS 364
Qy 384 RESDLALMTNAGTEIGVASTKAPTTLTLLMLVAKLSRLKGLDASIEHDIHVHGLQALPS 443
Db 365 LLTHCGVHINAGPEIGVASTKAYTSQFVAMVPALSISEDRAKQKQKREIMEGLAKVSE 424
Qy 444 RIBQMLSQDKRIEAL-AEDFSDXHHALFLSRGQOYPIALE 482
Db 425 QFKEILKNEPIKQKCAKFFKNQKSLLLGRCGQOFTALE 464

```

Search completed: June 14, 2006, 15:34:56
Job time : 22.6172 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:28:00 ; Search time 57.0938 Seconds
(without alignments)
933.660 Million cell updates/sec

Title: US-10-612-779-6

Perfect score: 3089

Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIKGTVDQPRNLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /EMC_Celestra_SID33/ptodata/2/iaa/5 COMB.pdp.*
- 2: /EMC_Celestra_SID33/ptodata/2/iaa/6 COMB.pdp.*
- 3: /EMC_Celestra_SID33/ptodata/2/iaa/7 COMB.pdp.*
- 4: /EMC_Celestra_SID33/ptodata/2/iaa/8 COMB.pdp.*
- 5: /EMC_Celestra_SID33/ptodata/2/iaa/9 COMB.pdp.*
- 6: /EMC_Celestra_SID33/ptodata/2/iaa/10 COMB.pdp.*
- 7: /EMC_Celestra_SID33/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3089	100.0	609	2	US-09-115-475-22
2	3072	99.4	609	2	US-09-115-475-28
3	3072	99.4	609	2	US-09-115-475-31
4	3068	99.3	609	2	US-09-115-475-16
5	3061	99.1	609	2	US-09-115-475-25
6	3053	98.8	609	2	US-09-115-475-19
7	2862	92.7	651	2	US-09-489-039A-10691
8	2498.5	80.9	622	2	US-09-543-681A-7522
9	1983	64.2	616	2	US-09-252-991A-19164
10	1769.5	57.3	613	2	US-09-328-352-5066
11	1491.5	48.3	619	2	US-09-540-236-2182
12	1351.5	43.8	572	2	US-09-302-540-16012
13	1346.5	43.6	592	1	US-08-599-171A-30
14	1346.5	43.6	592	1	US-08-646-590B-30
15	1346.5	43.6	592	2	US-09-069-226-30
16	1346.5	43.6	592	2	US-09-412-184-30
17	1148	37.2	611	2	US-09-438-185A-970
18	1132	36.6	655	2	US-09-107-532A-4425
19	1110	35.9	619	2	US-09-134-001C-4248
20	1106.5	35.8	602	2	US-09-583-110-4085
21	1106.5	35.8	635	2	US-09-107-433-5022
22	1098.5	35.6	682	1	US-08-911-445-3
23	1098.5	35.6	682	2	US-09-182-983-3
24	1098.5	35.6	682	2	US-09-771-838A-3
25	1098.5	35.6	682	2	US-09-949-016-6591
26	1078.5	34.9	680	2	US-09-731-166-6

27 955.5 30.9 615 1 US-08-911-445-2 Sequence 2, Appli
28 955.5 30.9 615 2 US-09-182-983-2 Sequence 2, Appli
29 955.5 30.9 615 2 US-09-771-838A-2 Sequence 2, Appli
30 839.5 27.2 640 2 US-09-134-000C-3871 Sequence 3871, Ap
31 739 23.9 487 2 US-09-248-796A-17250 Sequence 17250, A
32 698 22.6 365 2 US-09-949-016-11214 Sequence 232, App
33 670 21.7 405 2 US-09-710-279-232 Sequence 1214, A
34 605.5 19.6 422 2 US-09-270-767-45187 Sequence 45187, A
35 575 18.6 486 1 US-08-911-445-18 Sequence 18, Appl
36 575 18.6 486 2 US-09-182-983-18 Sequence 18, Appl
37 575 18.6 486 2 US-09-771-838A-18 Sequence 18, Appl
38 574 18.6 425 1 US-08-911-445-1 Sequence 1, Appli
39 574 18.6 425 2 US-09-182-983-1 Sequence 1, Appli
40 574 18.6 425 2 US-09-771-838A-1 Sequence 1, Appli
41 548 17.7 338 2 US-09-710-279-320 Sequence 320, App
42 495 16.0 307 2 US-09-198-452A-1041 Sequence 1041, Ap
43 360.5 11.7 161 2 US-09-198-452A-1040 Sequence 1040, Ap
44 326 10.6 708 2 US-09-252-991A-18074 Sequence 18074, A
45 314 10.2 182 2 US-09-198-452A-1042 Sequence 1042, Ap

ALIGNMENTS

RESULT 1

US-09-115-475-22

; Sequence-22, Application US/09115475

; Patent No. 6372457

; GENERAL INFORMATION:

; APPLICANT: Berry, Alan

; APPLICANT: Burlingame, Richard P.

; APPLICANT: Millis, James R.

; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE

; FILE REFERENCE: 3161-18-C1

; CURRENT APPLICATION NUMBER: US/09/115,475

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: PCT/US98/00800

; EARLIER FILING DATE: 1998-01-14

; EARLIER APPLICATION NUMBER: 60/035,494

; EARLIER FILING DATE: 1997-01-14

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-09-115-475-22

Query Match 100.0%; Score 3089; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 3.5e-300;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60
DB 1 MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60
QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPELREELKARGYTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMGENTFIASDQALLPVTRRFIFLEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMGENTFIASDQALLPVTRRFIFLEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRIHSGVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRIHSGVDLSLGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEITASEFRYKSAVRNRNSLMTILSQSGETADTLA 360

Db 301 CGTSYNSGMSRYWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

US-09-115-475-28
; Sequence 28, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-28

Query Match 99.4%; Score 3072; DB 2; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.8e-298;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTGEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLGKVMQAQAAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHNEPRLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHNEPRLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRISHGQVDLSGLGNADALLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRISHGQVDLSGLGNADALLSKVEHIQILA 300
QY 301 CGTSYNSGMSRYWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360

Db 301 CGTSYNSGMSRYWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3

US-09-115-475-31
; Sequence 31, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-31

Query Match 99.4%; Score 3072; DB 2; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.8e-298;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTGEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLGKVMQAQAAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPRLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPRLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
Db 181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRISHGQVDLSGLGNADALLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRISHGQVDLSGLGNADALLSKVEHIQILA 300
QY 301 CGTSYNSGMSRYWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360

Db 301 CGTSYNGMVSRYWFSFESLAGIPCDVEIASEFRYKSAVRNSLMTILTSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540
QY 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
Db 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 4

US-09-115-475-16
; Sequence 16, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-16

Query Match 99.3%; Score 3068; DB 2; Length 609;
Best Local Similarity 99.5%; Pred. No. 4.4e-298;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKQVOMLAQAAE 60
Db 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKQVOMLAQAAE 60
QY 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVEHIQILA 300
QY 301 CGTSYNGMVSRYWFSFESLAGIPCDVEIASEFRYKSAVRNSLMTILTSQSGETADTLA 360

Db 301 CGTSYNGMVSRYWFSFESLAGIPCDVEIASEFRYKSAVRNSLMTILTSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540
QY 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
Db 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 5

US-09-115-475-25
; Sequence 25, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-25

Query Match 99.1%; Score 3061; DB 2; Length 609;
Best Local Similarity 99.3%; Pred. No. 2.2e-297;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKQVOMLAQAAE 60
Db 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLGKQVOMLAQAAE 60
QY 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVEHIQILA 300
QY 301 CGTSYNGMVSRYWFSFESLAGIPCDVEIASEFRYKSAVRNSLMTILTSQSGETADTLA 360

```
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGGQOYPTA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGGQOYPTA 480
QY 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAQSVTVE 609
Db 601 NLAQSVTVE 609
```

RESULT 6

```
US-09-115-475-19
; Sequence 19, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Mills, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-19
```

```
Query Match 98.8%; Score 3053; DB 2; Length 609;
Best Local Similarity 99.0%; Pred. No. 1.4e-296;
Matches 603; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRTSHGQVDLSLSELPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRTSHGQVDLSLSELPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
```

```
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGGQOYPTA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGGQOYPTA 480
QY 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAQSVTVE 609
Db 601 NLAQSVTVE 609
```

RESULT 7

```
US-09-489-039A-10691
; Sequence 10691, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10691
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10691
```

```
Query Match 92.7%; Score 2862; DB 2; Length 651;
Best Local Similarity 91.8%; Pred. No. 2.2e-277;
Matches 559; Conservative 29; Mismatches 21; Indels 0; Gaps 0;
```

```
QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 60
Db 43 MCGIVGAIAQRDVAEILLEGRLRLEYGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 102
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
Db 103 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 162
QY 121 SETDTEVIAHLVNNWELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
Db 163 TETDTEVIAHLVHWELEQGGTLREAVLRAIPQLRGAYGTVIMDTPDGTLLAARSGSPV 222
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 223 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKSAQVRKQETESNLQ 282
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRTSHGQVDLSLSELPNADLLSKVEHIQILA 300
Db 283 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRTSHGQVDLSLSELPNADLLSKVEHIQILA 342
QY 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Db 343 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 402
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQTLTVLLMLVAKL 420
```

```
Db 403 RLKELGSLGSLALCNVPGSSLVRESLALMTKAGTEIGVASTKFTTQLTVLLMLVAKL 462
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEDFSXKHALFLSRGDQYPIA 480
Db 463 ARUKGODASIEHDIVHGLQALPNRIEQMLSDKRIEQLAEFRFSXKHALFLSRGDQYPIA 522
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLUKSNIEEVRAR 540
Db 523 MEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLUKSNIEEVRAR 582
Qy 541 GGOLYVPADQADGAFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
Db 583 GGELYVPADGEGAGFSGDNNHIIEMPHVEETIAPIFTVPLQLLAYHVALIKGTDVDQPR 642
Qy 601 NLAKSVTVE 609
Db 643 NLAKSVTVE 651

RESULT 8
US-09-543-681A-7522
; Sequence 7522, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7522
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7522

Query Match 80.9%; Score 2498.5; DB 2; Length 622;
Best Local Similarity 80.8%; Pred. No. 5.6e-241;
Matches 492; Conservative 46; Mismatches 70; Indels 1; Gaps 1;

Qy 1 MCGIVGAIAORDVAEILLEGRLRLRYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
Db 15 MCGIVGAIAORDVAEILLEGRLRLRYGYDSAGLAVVDNDCHMTRRLREVGVQMLABEAE 74
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 75 KTVQIGGTGTAHTRWATHGEPSEVNAHPHYSGTIAVHNGIIEHHEPLREELKARGYTFV 134
Qy 121 SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 135 SQTDTTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 194
Qy 181 IGLWGENFIASDQALALLPVTTRFIFLEBGDIATRISVNIIFDKTGAEVKRODIESNLQ 240
Db 195 VGLWGENFIASDQALALLPVTTRFIFLEBGDIATRISVNIIFDKTGAEVKRODIESNLQ 254
Qy 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Db 255 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 314
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNNSLMTLSQSGETADTLAAGL 360
Db 315 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNNSLMTLSQSGETADTLAAGL 374
Qy 361 RLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 375 RLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 434
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEDFSXKHALFLSRGDQYPIA 480
```

```
Db 435 GRKGVEA-LEQQVAMALHALPSRIESMLSKDVIEALAEDFSXKHALFLSRGDQYPIA 493
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLUKSNIEEVRAR 540
Db 494 VEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLUKSNIEEVRAR 553
Qy 541 GGOLYVPADQADGAFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
Db 554 GGELYVPADQADGAFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 613
Qy 601 NLAKSVTVE 609
Db 614 NLAKSVTVE 622

RESULT 9
US-09-252-991A-19164
; Sequence 19164, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; CURRENT APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19164
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19164

Query Match 64.2%; Score 1983; DB 2; Length 616;
Best Local Similarity 62.4%; Pred. No. 2.6e-189;
Matches 381; Conservative 101; Mismatches 127; Indels 2; Gaps 2;

Qy 1 MCGIVGAIAORDVAEILLEGRLRLRYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
Db 6 MCGIVGAIAERNTPILIEGLKRLRYGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 65
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHY-SEHIVVHNGIIEHHEPLREELKARGYTF 119
Db 66 GTPELLGRGLGIAHTRWATHGAPTEGNAHPHFSSDELAHVHNGIIEHHEPLREELKARGYTF 125
Qy 120 VSETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPL 179
Db 126 TSQTDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPL 185
Qy 180 VIGLWGENFIASDQALALLPVTTRFIFLEBGDIATRISVNIIFDKTGAEVKRODIESNL 239
Db 186 VIGLWGENFIASDQALALLPVTTRFIFLEBGDIATRISVNIIFDKTGAEVKRODIESNL 245
Qy 240 QYDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQIL 299
Db 246 GAERADKGEYRHFMKLEIHEQPSVQVQRTLEGRLQGNQVMVESFGPQAAELFAKVRNQIV 305
Qy 300 ACCTSNGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNNSLMTLSQSGETADTLAAGL 359
Db 306 ACCTSNGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNNSLMTLSQSGETADTLAAGL 365
Qy 360 LRLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 419
Db 366 LRLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 425
Qy 420 LRLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 478
Db 426 LRLKELGSLGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 485
```


THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:15:09 ; Search time 305.293 Seconds
(without alignments)
1845.226 Million cell updates/sec

Title: US-10-612-779-6
Perfect score: 3089
Sequence: 1 MCGIVGAIQRDAEILLEG.....LIKGTVDQPRNLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	3068	99.3	609	2	Q2M847_ECOLI	Q2M847 escherichia
2	3063	99.2	608	1	GLMS_ECOLI	P17169 e glucosami
3	3063	99.2	608	1	GLMS_ECOLI	P17169 e glucosami
4	3061	99.1	609	2	Q329R8_SHIDS	Q329R8 shigella dy
5	3059	99.0	608	1	GLMS_ECOLI	Q310M9 shigella bo
6	3059	99.0	608	1	GLMS_SHIFL	Q8xeg2 e glucosami
7	3055	98.9	609	2	Q3YVW3_SHIDS	Q83iy4 s glucosami
8	3049	98.7	608	1	GLMS_ECOLI	Q3YVW3 shigella so
9	3040	98.4	608	1	GLMS_SALPA	Q8fbt4 e glucosami
10	3040	98.4	609	2	Q57HY2_SALCH	Q5pkv9 s glucosami
11	3037	98.3	608	1	GLMS_SALTY	Q57HY2 salmonella
12	3030	98.1	608	1	GLMS_SALTY	Q8zkl1 s glucosami
13	2667	86.4	609	1	GLMS_ERACT	Q8z2q2 s glucosami
14	2663	86.2	608	1	GLMS_YERPE	Q8cyj9 e glucosami
15	2663	86.2	608	1	GLMS_YERPS	Q8z9s8 y glucosami
16	2553	82.6	608	1	GLMS_PHOLL	Q663r1 y glucosami
17	2532	82.0	610	2	Q2N083_SODGL	Q7na97 p glucosami
18	2260	73.2	610	2	Q4QNB6_HAE18	Q2nq83 sodalis glo
19	2251	72.9	609	1	GLMS_HAEIN	Q4qnb6 haemophilus
20	2245	72.7	609	1	GLMS_VIBCH	Q9kun8 h glucosami
21	2243	72.6	609	1	GLMS_VIBF1	Q8z279 v glucosami
22	2236	72.4	609	1	GLMS_PHOPR	Q61l13 p glucosami
23	2214	71.7	609	1	GLMS_VIBPA	Q87sr3 v glucosami
24	2210	71.6	609	1	GLMS_VIBVU	Q8def3 v glucosami
25	2208	71.5	609	1	GLMS_VIBVY	Q7mp62 v glucosami
26	2200	71.2	610	2	Q3EGX5_ACTSC	Q3egx5 actinobacil
27	2193	70.7	609	1	GLMS_PASMU	P57963 p glucosami
28	2166	70.1	610	2	Q47UE2_COLP3	Q47ue2 colwellia p
29	2145	69.4	609	2	Q338I7_9GAMM	Q33si7 shewanella
30	2145	69.4	609	2	Q3P1V9_9GAMM	Q3plv9 shewanella
31	2142	69.3	609	2	Q3Q819_9GAMM	Q3q819 shewanella

32	2140	69.3	609	2	Q2X1G5_9GAMM	Q2x1g5 shewanella
33	2140	69.3	609	2	Q2ZQ11_SHEPU	Q2zq11 shewanella
34	2135	69.1	609	2	Q368R1_9GAMM	Q368r1 shewanella
35	2134	69.1	609	2	Q2Z4D9_9GAMM	Q2z4d9 shewanella
36	2132	69.0	609	2	Q35V59_9GAMM	Q35v59 shewanella
37	2126	68.8	609	1	GLMS_HAE18	Q7vkk4 h glucosami
38	2126	68.8	608	1	GLMS_SHEON	Q8cx33 s glucosami
39	2125	68.8	609	2	Q3NXY7_SHEPR	Q3nxy7 shewanella
40	2099	68.0	609	1	GLMS_IDILO	Q5qz15 i glucosami
41	2088	67.6	610	2	Q3CSC7_ALTAT	Q3csc7 pseudalter
42	2000	64.8	606	2	Q3IK37_PSEHT	Q3ik37 pseudalter
43	1992	64.5	611	2	Q2XHY7_PSEPU	Q2xhy7 pseudomonas
44	1990	64.4	610	2	Q3K445_PSEPF	Q3k445 pseudomonas
45	1988	64.4	610	1	GLMS_PSEPK	Q88bx8 p glucosami

ALIGNMENTS

RESULT 1
Q2M847_ECOLI
ID Q2M847_ECOLI PRELIMINARY; PRT; 609 AA.
AC Q2M847;
DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, sequence version 1.
DT 07-MAR-2006, entry version 2.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=glms;
OS Escherichia coli W3110.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=316407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=810531692; PubMed=6159575;
RA Smith D.R., Calvo J.M.;
RT "Nucleotide sequence of the E coli gene coding for dihydrofolate reductase.";
RL Nucleic Acids Res. 8:2255-2274(1980).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Sakiguchi T., Ortega-Cesena J., Nosoh Y., Ohashi S., Tsuda K., Kanaya S.;
RT "DNA and amino-acid sequences of 3-isopropylmalate dehydrogenase of RT Bacillus coagulans. Comparison with the enzymes of Saccharomyces cerevisiae and Thermus thermophilus.";
RL Biochim. Biophys. Acta 867:36-44(1986).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Chen H., Sun Y., Stark T., Beattie W., Moses R.E.;
RT "Nucleotide sequence and deletion analysis of the polB gene of Escherichia coli.";
RL DNA Cell Biol. 9:613-635(1990).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Snallegah J.E., Kelln R.A.;
RT "Cloning, nucleotide sequence and expression of the Escherichia coli K-12 pyrH gene encoding UMP kinase.";
RL Genetics (Life Sci. Adv.) 11:59-65(1992).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Hayaishi K., Morooka N., Yamamoto Y., Fujita K., Isono K., Choi S., Ohtsubo E., Baba T., Wanner B.L., Mori H., Horiuchi T.;
RT "Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110.";
RL Mol. Syst. Biol. 0:0-0(2006).
RN [6]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=K-12;
RX PubMed=16397293; DOI=10.1093/nar/gkj150;
RA Riley M., Abe T., Arnaud M.B., Berlyn M.K., Blattner F.R.,
RA Chaudhuri R.R., Glasner J.D., Horiuchi T., Kessler I.M., Kosuge T.,
RA Mori H., Perna N.T., Plunkett G. III, Rudd K.E., Serres M.H.,
RA Thomas G.H., Thomson N.R., Wishart D., Wanner B.L.;
RT "Escherichia coli K-12: a cooperatively developed annotation snapshot-
RT -2005";
RL Nucleic Acids Res. 34:1-9 (2006).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97349980; PubMed=9205837; DOI=10.1093/dnares/4.2.91;
RA Yamanoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y.,
RA Oshima T., Oyama S., Saito N., Sampaio G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113 (1997).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251358; PubMed=9097040; DOI=10.1093/dnares/3.6.379;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampaio G., Seki Y.,
RA Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamanoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map";
RL DNA Res. 3:379-392 (1996).
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251357; PubMed=9097039; DOI=10.1093/dnares/3.6.363;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampaio G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map";
RL DNA Res. 3:363-377 (1996).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97094878; PubMed=8940112; DOI=10.1074/jbc.271.49.31145;
RA Arn E.A., Abelson J.N.;
RT "The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and
RT genomic disruption";
RL J. Biol. Chem. 271:31145-31153 (1996).
RN [11]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampaio G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map";
RL DNA Res. 3:137-155 (1996).
RN [12]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;

RC STRAIN=K-12;
RX MEDLINE=94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 2.4-4.1 min (110,917-193,643 bp) region";
RL Nucleic Acids Res. 22:1637-1639 (1994).
RN [13]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94240115; PubMed=8183897;
RA Janosi L., Shimizu I., Kaji A.;
RT "Ribosome recycling factor (ribosome releasing factor) is essential
RT for bacterial growth";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4249-4253 (1994).
RN [14]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94124004; PubMed=7904973; DOI=10.1016/0378-1119(93)90470-N;
RA Allikmets R., Gerrard B.C., Court D., Dean M.C.;
RT "Cloning and organization of the abc and mdl genes of Escherichia
RT coli: relationship to eukaryotic multidrug resistance";
RL Gene 136:231-236 (1993).
RN [15]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94018640; PubMed=8412694;
RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.D.;
RT "The genes of the glutamine synthetase adenylation cascade are not
RT regulated by nitrogen in Escherichia coli";
RL Mol. Microbiol. 9:443-458 (1993).
RN [16]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93259920; PubMed=8387990;
RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
RT "Rho elements of Escherichia coli K-12: complex composites of shared
RT and unique components that have different evolutionary histories";
RL J. Bacteriol. 175:2799-2808 (1993).
RN [17]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93121180; PubMed=8419307;
RA Yamada M., Asaka S., Saier M.H. Jr., Yamada Y.;
RT "Characterization of the gcd gene from Escherichia coli K-12 W3110 and
RT regulation of its expression";
RL J. Bacteriol. 175:568-571 (1993).
RN [18]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93116053; PubMed=1474579;
RA Cormack R.S., Mackie G.A.;
RT "Structural requirements for the processing of Escherichia coli 5 S
RT ribosomal RNA by RNase E in vitro";
RL J. Mol. Biol. 228:1078-1090 (1992).
RN [19]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93094132; PubMed=1459951;
RA Gervais F.G., Drapeau G.R.;
RT "Identification, cloning, and characterization of rcsF, a new
RT regulator gene for exopolysaccharide synthesis that suppresses the
RT division mutation ftsZ84 in Escherichia coli K-12";
RL J. Bacteriol. 174:8016-8022 (1992).
RN [20]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93077430; PubMed=1447125;
RA Yamanaka K., Ogura T., Niki H., Hiraga S.;
RT "Identification and characterization of the smbA gene, a suppressor of
RT the mukB null mutant of Escherichia coli";
RL J. Bacteriol. 174:7517-7526 (1992).
RN [21]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;

```
RX MEDLINE=93011013; PubMed=1396599;
RA Condon C., Phillips J., Fu Z.Y., Squires C., Squires C.L.;
Query Match 99.3%; Score 3068; DB 2; Length 609;
Best Local Similarity 99.5%; Pred. No. 7.8e-185;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYSAGLAVVDTECHMTLRRLRGKVMQLAQAAE 60
DB 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYSAGLAVVDTECHMTLRRLRGKVMQLAQAAE 60
QY 61 EHLPGGTGTAHTRWATHGSPSEVNAHPHYSEHIVVHNGIIEHNEPRLRELKARGYTFV 120
DB 61 EHLPGGTGTAHTRWATHGSPSEVNAHPHYSEHIVVHNGIIEHNEPRLRELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPV 180
DB 121 SETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPV 180
QY 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSYNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSYNIFDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
QY 301 CGTSYNGMSRVYWFESLAGIPCDEIASEFRYKSAVRNRLMITLSQSGETADTLA 360
DB 301 CGTSYNGMSRVYWFESLAGIPCDEIASEFRYKSAVRNRLMITLSQSGETADTLA 360
QY 361 RLSKELGYLSLAI CNVPGSSLYRESLALMTNAGTIGVASTKFTOLTVLMLVAKL 420
DB 361 RLSKELGYLSLAI CNVPGSSLYRESLALMTNAGTIGVASTKFTOLTVLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQWLSODKRIEALAEFSDKHALFLSRGDQYPIA 480
DB 421 SRLKGLDASIEHDI VHGLQALPSRIEQWLSODKRIEALAEFSDKHALFLSRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVPADODAGFVSSDNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
DB 541 GGQLYVPADODAGFVSSDNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609
RESULT 2
GLMS_ECOLI STANDARD; PRT; 608 AA.
AC P17169; P17645;
DT 01-AUG-1990, integrated into UniProtKB/Swiss-Prot.
DT 15-DEC-1998, sequence version 3.
DT 07-MAR-2006, entry version 70.
DE Glucosamine-6-phosphate aminotransferase (isomerizing)
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=glms; Ordered(LocusNames=b3729;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=85121806; PubMed=6395859;
RA Walker J.E., Gay N.J., Saraste M., Eberle A.N.;
RT "DNA sequence around the Escherichia coli unc operon. Completion of
the sequence of a 17 kilobase segment containing asnA, oriC, unc, glms
and phoS."; 224:799-815 (1984).
[2]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=93315143; PubMed=7686982;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication.";
Genomics 16:551-561 (1993).
[3]
RN PROTEIN SEQUENCE OF 48-51; 218-230; 488-492; 504-507 AND 600-608.
RX MEDLINE=92007872; PubMed=1915361;
RA Gollineili-Pimpaneau B., Badet B.;
RT "Possible involvement of lys603 from Escherichia coli glucosamine-6-
RT phosphate synthase in the binding of its substrate fructose 6-
phosphate.";
RL Eur. J. Biochem. 201:175-182 (1991).
[4]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 489-608.
RX MEDLINE=82200022; PubMed=6283361;
RA Lichtenstein C., Bremner S.;
RT "Unique insertion site of Tn7 in the E. coli chromosome.";
RL Nature 297:601-603 (1982).
[5]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 606-608.
RX MEDLINE=86215091; PubMed=3010949;
RA Gay N.J., Tybulewicz V.L.J., Walker J.E.;
RT "Insertion of transposon Tn7 into the Escherichia coli glms
transcriptional terminator.";
RL Biochem. J. 234:111-117 (1986).
[6]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 596-608.
RX MEDLINE=88086894; PubMed=2826397;
RA McKown R.L., Orle K.A., Chen T., Craig N.L.;
RT "Sequence requirements of Escherichia coli attTn7, a specific site of
transposon Tn7 insertion.";
RL J. Bacteriol. 170:352-358 (1988).
[7]
RN CHARACTERIZATION.
RX MEDLINE=88281539; PubMed=3134953; DOI=10.1016/0300-9084(88)90073-9;
RA Dutka-Mallen S., Mazodier P., Badet B.;
RT "Molecular cloning and overexpression of the glucosamine synthetase
gene from Escherichia coli.";
RL Biochimie 70:287-290 (1988).
[8]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-240.
RX MEDLINE=96434326; PubMed=8805567; DOI=10.1016/S0969-2126(96)00087-1;
RA Isupov M.N., Obmolova G., Butterworth S., Badet-Denisot M.-A.,
RA Badet B., Polikarpov I., Littlechild J.A., Teplyakov A.;
RT "Substrate binding is required for assembly of the active conformation
of the catalytic site in Ntn amidotransferases: evidence from the 1.8-
RT A crystal structure of the glutaminase domain of glucosamine 6-
phosphate synthase.";
RL Structure 4:801-810 (1996).
[9]
RN X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 243-608.
RX MEDLINE=98416699; PubMed=9739095; DOI=10.1016/S0969-2126(98)00105-1;
RA Teplyakov A., Obmolova G., Badet-Denisot M.-A., Badet B.,
RA Polikarpov I.;
RT "Involvement of the C terminus in intramolecular nitrogen channeling
in glucosamine 6-phosphate synthase: evidence from a 1.6-A crystal
RT structure of the isomerase domain.";
RL Structure 6:1047-1055 (1998).
[10]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 241-608.
RX MEDLINE=99190083; PubMed=10091662;
RA Teplyakov A., Obmolova G., Badet-Denisot M.A., Badet B.;
RT "The mechanism of sugar phosphate isomerization by glucosamine 6-
RT phosphate synthase.";
RL Protein Sci. 8:596-602 (1999).
-1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
```

converting fructose-6P into glucosamine-6P using glutamine as a nitrogen source.
 -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-glutamate + D-glucosamine 6-phosphate.
 -!- SUBUNIT: Homodimer.
 -!- INTERACTION:
 P0957:ala8; NbExp=1; IntAct=EBI-551022, EBI-544061;
 P08622:dnaj; NbExp=1; IntAct=EBI-551022, EBI-545285;
 P76552:euth; NbExp=1; IntAct=EBI-551022, EBI-551031;
 P62615:ispe; NbExp=1; IntAct=EBI-551022, EBI-562202;
 P61175:rpiv; NbExp=1; IntAct=EBI-551022, EBI-542255;
 P21166:trkh; NbExp=1; IntAct=EBI-551022, EBI-550268;
 P76093:ymbd; NbExp=1; IntAct=EBI-551022, EBI-551038;
 P33366:yohd; NbExp=1; IntAct=EBI-551022, EBI-551046;
 -!- SUBCELLULAR LOCATION: Cytoplasm.
 -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
 -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
 -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
 Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 Distributed under the Creative Commons Attribution-NoDerivs License

 EMBL: X01631; CAA25785.1; -; Genomic DNA.
 EMBL: L10328; AAG62080.1; -; Genomic DNA.
 EMBL: U00096; AAC76752.1; -; Genomic DNA.
 EMBL: V00620; CAA23894.1; -; Genomic DNA.
 EMBL: M18980; AAA23836.1; -; Genomic DNA.
 EMBL: B65176; XNECGM.
 PIR: B65176; X-ray; A/B/C=1-608.
 PDB: 1JXA; X-ray; A/B/C=1-608.
 PDB: 1MOQ; X-ray; @=241-608.
 PDB: 1MOR; X-ray; @=241-608.
 PDB: 1MOS; X-ray; A=241-608.
 PDB: 1XFF; X-ray; A/B=1-240.
 PDB: 1XFG; X-ray; A/B=1-240.
 PDB: 2BFJ; X-ray; A/B=1-608.
 PDB: 2BPL; X-ray; A/B/C=1-608.
 MEROPS: C44.971; -.
 Genomereviews: U00096_GR; b3729.
 EcoBASE; EB0377; -.
 EcoGene; EG10382; glms.
 BioCyc; EcoCyc: L-GLN-FRUCT-6-P-AMINOTRANS-MONOMER; -.
 GO; GO:0005515; F:protein binding; IPI.
 HAMAP; MF_00164; -; 1.
 InterPro; IPR000583; GATase_2.
 InterPro; IPR005855; Glms_trans.
 InterPro; IPR001347; SIS.
 Pfam; PF00310; GATase 2; 1.
 Pfam; PF01380; SIS; 2.
 TIGRFAMs; TIGR01135; glms; 1.
 PROSITE; P500443; GATASE_TYPE_II; 1.
 3D-structure; Amino transferase; Complete proteome;
 Direct protein sequencing; Glutamine amidotransferase; Transferase.
 INIT MET 0 0
 CHAIN 1 608 Glucosamine--fructose-6-phosphate
 aminotransferase [isomerizing].
 /FTid=PRO_0000135328.
 Glutamine amidotransferase.
 GATase.
 Isomerization Fru-6P.
 KL -> NV (in Ref. 1).
 REGION 1 181
 ACT_SITE 1 1
 ACT_SITE 603 603
 CONFLICT 418 419
 STRAND 2 7
 STRAND 9 10
 STRAND 13 26
 STRAND 29 36
 STRAND 38 39
 STRAND 42 49
 STRAND 51 60
 STRAND 61 61
 STRAND 66 73
 STRAND 76 78
 STRAND 80 81
 STRAND 82 84
 STRAND 85 86

FT STRAND 88 90
 FT TURN 91 92
 FT STRAND 93 101
 FT TURN 102 103
 FT HELIX 104 113
 FT TURN 114 115
 FT STRAND 120 120
 FT STRAND 123 123
 FT STRAND 124 136
 FT TURN 137 137
 FT STRAND 138 139
 FT HELIX 141 151
 FT STRAND 154 162
 FT TURN 163 164
 FT TURN 166 167
 FT STRAND 169 176
 FT STRAND 179 182
 FT STRAND 184 185
 FT STRAND 187 192
 FT HELIX 193 195
 FT TURN 196 199
 FT STRAND 201 205
 FT TURN 208 209
 FT STRAND 211 214
 FT STRAND 216 217
 FT STRAND 219 222
 FT TURN 224 225
 FT STRAND 228 229
 FT STRAND 233 235
 FT TURN 246 247
 FT STRAND 249 250
 FT HELIX 251 257
 FT TURN 258 258
 FT HELIX 259 267
 FT TURN 268 269
 Query Match 99.2%; Score 3063; DB 1; Length 608;
 Best Local Similarity 99.5%; Pred.No.1.6e-184; Indels 0; Gaps 0;
 Matches 605; Conservative 0; Mismatches 3;
 QY 2 CGIVGAIQAQDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKVQMLQAASE 61
 DB 1 CGIVGAIQAQDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRLRGKVQMLQAASE 60
 QY 62 HPLHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFVS 121
 DB 61 HPLHGGTGAIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFVS 120
 QY 122 ETDTEVIAHLVNWELKQGGTTLREAVLRAIPOLRGAYGTIVIMDSRHPDITLLAARSGSPLVI 181
 DB 121 ETDTEVIAHLVNWELKQGGTTLREAVLRAIPOLRGAYGTIVIMDSRHPDITLLAARSGSPLVI 180
 QY 182 GLGMGENFIASDQALLPVTTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQY 241
 DB 181 GLGMGENFIASDQALLPVTTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQY 240
 QY 242 DAGDKGIYCHVMQKEIYEQNAIKNTLTGRISHGQVDLSELGNADLLSKVEHTQILAC 301
 DB 241 DAGDKGIYCHVMQKEIYEQNAIKNTLTGRISHGQVDLSELGNADLLSKVEHTQILAC 300
 QY 302 GTSYNSGMVSRVWFESLAGICPDVEIASEFFRYKSAVRNSLMTLSOSGETADTLAIGLR 361
 DB 301 GTSYNSGMVSRVWFESLAGICPDVEIASEFFRYKSAVRNSLMTLSOSGETADTLAIGLR 360
 QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 421
 DB 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 420
 QY 422 RLKGLDASIEHDIHVLGLOALPSRIEOMLSQDKRIEALAEFSDKHALFLSGDQYPIAL 481
 DB 421 RLKGLDASIEHDIHVLGLOALPSRIEOMLSQDKRIEALAEFSDKHALFLSGDQYPIAL 480
 QY 482 EGALKLKEISVTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAAG 541

```

Db 481 EGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 540
Qy 542 GOLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDPNRN 601
Db 541 GOLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDPNRN 600
Qy 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 3
Q329R8_SHIDS
ID Q329R8_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q329R8
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name-glmS; OrderedLocusNames=SDY_4019;
OS Shigella dysenteriae serotype 1 (strain Sdl97).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery.";
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000034; ABB63937.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . ; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66904 MW; 6373301513227484 CRC64;

Query Match 99.2%; Score 3063; DB 2; Length 609;
Best Local Similarity 99.3%; Pred. No. 1.6e-184;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQDVAAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60
Db 1 MCGIVGAIAQDVAAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHVNWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHVNWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

Qy 181 IGLGMGNFTASDQALLPYTRFIFUEEGDIAETRRSVNIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGNFTASDQALLPYTRFIFUEEGDIAETRRSVNIPDKTGAEVKRODIESNLQ 240

Qy 241 YDAGDKGIYCHYMOKETIYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKETIYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

```

```

Qy 301 CGTSYNSGMVSRWFESLAGIPCDVETASEFRYKRSVAVRNSLWITLSQSGETADTLAAGL 360
Db 301 CGTSYNSGMVSRWFESLAGIPCDVETASEFRYKRSVAVRNSLWITLSQSGETADTLAAGL 360
Qy 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFVTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFVTTQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDPNR 600
Db 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDPNR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 4
Q31UM9_SHIDS
ID Q31UM9_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q31UM9
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name-glmS; OrderedLocusNames=SBO_3758;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300268;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery.";
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000036; ABB68229.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . ; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66898 MW; 989B94DC79C348D4 CRC64;

Query Match 99.1%; Score 3061; DB 2; Length 609;
Best Local Similarity 99.2%; Pred. No. 2.2e-184;
Matches 604; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQDVAAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60
Db 1 MCGIVGAIAQDVAAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKGVOMLAQAAE 60

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHEPLREELKARGYTFV 120

```

Db	61	EHPLHGGTGIAHTRWATHGEPSEANAHVSHIMVWHNGIIEHPEUREELKARGYTFV	120
Qy	121	SETDTEVIAHLVNWELKQGTLEAREAVLRAIPQLRGAYGVIMDSRHPDTLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNWELKQGTLEAREAVLRAIPQLRGAYGVIMDSRHPDTLLAARSGSPLV	180
Qy	181	IGLGMENFIASDQALLPVTFRFFLEBGDIAETRRSVNIFDKTGAEVRKQDIESNLQ	240
Db	181	IGLGMENFIASDQALLPVTFRFFLEBGDIAETRRSVNIFDKTGAEVRKQDIESNLQ	240
Qy	241	YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA	300
Qy	301	CGTSYNSGMVSRYPFESLAGICDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL	360
Db	301	CGTSYNSGMVSRYPFESLAGICDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL	360
Qy	361	RLSKELGYLGSIAICNPVSSIVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL	420
Db	361	RLSKELGYLGSIAICNPVSSIVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL	420
Qy	421	SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHLFLSRGDOYPTA	480
Db	421	SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHLFLSRGDOYPTA	480
Qy	481	LEGALKLKEISYIHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR	540
Db	481	LEGALKLKEISYIHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR	540
Qy	541	GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR	600
Db	541	GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLQLLAYHVALIKGTVDQPR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	
RESULT 5			
GLMS ECO57			
ID	GLMS ECO57	STANDARD;	PRT; 608 AA.
AC	Q8XEG2;		
DT	11-JUL-2002,	integrated into UniProtKB/Swiss-Prot.	
DT	11-JUL-2002,	sequence version 2.	
DT	07-MAR-2006,	entry version 24.	
DE	Glucosamine-6-phosphate	aminotransferase [isomerizing]	
DE	(EC 2.6.1.16)	(Hexosephosphate aminotransferase (D-fructose-6-phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate	
DE	Name=glms; OrderedLocustNames=25227, ECs4671;		
GN	Escherichia coli O157:H7.		
OS	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=83334;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;		
RX	MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;		
RA	Perna N.T., Plunkett G., Evans P.S., Gregor J., Kirkpatrick H.A.,		
RA	Rose D.J., Mayhew G.F., Klink S., Boutin A., Shao Y., Miller L.,		
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,		
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,		
RA	Welch R.A., Blattner F.R.;		
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"		
RL	Nature 409:529-533 (2001).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;		
RX	MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;		
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,		

RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,		
RA	Lida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,		
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.;		
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli		
RT	O157:H7 and genomic comparison with a laboratory strain K-12.;"		
RL	DNA Res. 8:11-22(2001).		
CC	-!- FUNCTION: Catalyzes the first step in hexosamine metabolism,		
CC	converting fructose-6P into glucosamine-6P using glutamine as a		
CC	nitrogen source (By similarity).		
CC	-!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-		
CC	glutamate + D-glucosamine 6-phosphate.		
CC	-!- SUBUNIT: Homodimer (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).		
CC	-!- SIMILARITY: In the C-terminal section; belongs to the SIS family.		
CC	GFAT subfamily.		
CC	-!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.		
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms		
CC	Distributed under the Creative Commons Attribution-NoDerivs License		
CC	EMBL; AB005174; AAG58932.1; -; Genomic DNA.		
DR	EMBL; BA000007; BAB38094.1; -; Genomic DNA.		
DR	PIR; G91212; G91212.		
DR	PIR; H86058; H86058.		
DR	HSP; P17169; 1JXA.		
DR	SMR; Q8XEG2; 1-239, 243-608.		
DR	GenomeReviews; BA000007 GR; ECs4671.		
DR	GenomeReviews; AB005174 GR; z5227.		
DR	BioCyc; ECOL8334-1; ECs4671-MONOMER; -.		
DR	HMAP; MF 00164; -; 1.		
DR	InterPro; IPR000583; GATase_2.		
DR	InterPro; IPR005855; GlnS_trans.		
DR	InterPro; IPR001347; SIS.		
DR	Pfam; PF00310; GATase_2; 1.		
DR	Pfam; PF01380; SIS; 2.		
DR	TIGRFAMS; TIGR01135; glms; 1.		
DR	PROSITE; PS00443; GATASE_TYPE_II; 1.		
KW	Aminotransferase; Complete proteome; Glutamine amidotransferase;		
KW	Transferase.		
FT	INIT MET 0 0		
FT	CHAIN 1 608		
FT	By similarity.		
FT	Glucosamine--fructose-6-phosphate		
FT	aminotransferase [isomerizing].		
FT	/FTid=PRO 0000135330.		
FT	Glutamine amidotransferase.		
FT	REGION 1 240		
FT	ACT SITE 1 1		
FT	ACT SITE 603 603		
FT	CONFLICT 566 566		
FT	H -> N (in Ref. 2).		
SQ	SEQUENCE 608 AA; 66777 MW; 882EDA38B7F67148 CRC64;		
Query Match 99.0%; Score 3059; DB 1; Length 608;			
Best Local Similarity 99.3%; Pred.No. 2.9e-184;			
Matches 604; Conservative 1; Mismatches 3; Indels 0; Gaps 0;			
Qy	2	CGIVGAIAQRDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLKQVQLAAQAAE	61
Db	1	CGIVGAIAQRDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLKQVQLAAQAAE	60
Qy	62	HPLHGGTGIAHTRWATHGEPSEVNAHPHSEHIVVWHNGIIEHPEUREELKARGYTFVS	121
Db	61	HPLHGGTGIAHTRWATHGEPSEVNAHPHSEHIVVWHNGIIEHPEUREELKARGYTFVS	120
Qy	122	ETDTEVIAHLVNWELKQGTLEAREAVLRAIPQLRGAYGVIMDSRHPDTLLAARSGSPLVI	181
Db	121	ETDTEVIAHLVNWELKQGTLEAREAVLRAIPQLRGAYGVIMDSRHPDTLLAARSGSPLVI	180
Qy	182	GLGMENFIASDQALLPVTFRFFLEBGDIAETRRSVNIFDKTGAEVRKQDIESNLQ	241
Db	181	GLGMENFIASDQALLPVTFRFFLEBGDIAETRRSVNIFDKTGAEVRKQDIESNLQ	240
Qy	242	DAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC	301
Db	241	DAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC	300


```

QY 302 GTSYNSGMVSKYWPESLAGIPCDVEIASERFYSKSAVRNSLMTLSQSGETADTLAQLR 361
Db 301 GTSYNSGMVSKYWPESLAGIPCDVEIASERFYSKSAVRNSLMTLSQSGETADTLAQLR 360
QY 362 LSKELGYGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLVLLMLVAKLS 421
Db 361 LSKELGYGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSRGQDYPAL 481
Db 421 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSRGQDYPAL 480
QY 482 EGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 541
Db 481 EGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPITFTVPLQLLAYHVALIKGTDVDDQPRN 601
Db 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPITFTVPLQLLAYHVALIKGTDVDDQPRN 600
QY 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 6
GLMS_SHIFL STANDARD; PRT; 608 AA.
AC Q831Y4; Q7BZ99;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 22.
DE Glucosamine--fructose-6-phosphate aminotransferase (isomerizing)
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate aminotransferase) (G6PAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Names: glms; OrderedLocusNames=SF3809, S3959;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a; DOI=10.1093/nar/gkf566;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.P., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -I- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -I- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -I- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -I- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC G6PAT subfamily.

```

```

CC -I- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC -----
CC ENBL; AE005674; AAN45249.1; -; Genomic DNA.
CC ENBL; AE016991; AAP18948.1; -; Genomic DNA.
CC HSSP; P17169; 1JXA.
CC SNR; Q831Y4; 1-239, 243-608.
CC GenomeReviews; AE014073_GR; S3959.
CC GenomeReviews; AE005674_GR; SF3809.
CC BioCyc; SFLE198214:AAN45249.1-MONOMER; -.
CC HAMAP; MF 00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005885; Glms_trans.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC Pfam; PF01380; SIS; 2.
CC TIGRFAMs; TIGR01135; glms; 1.
CC PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT CHAIN 0 0 By similarity.
FT INIT MET 1 608 Glucosamine--fructose-6-phosphate
FT FT aminotransferase [isomerizing].
FT FT /FTID-PRO 0000135377,
FT FT Glutamine amidotransferase.
FT REGION 1 240
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66735 MW; B0CEDA38B6F00F7D CRC64;

Query Match 99.0%; Score 3059; DB 1; Length 608;
Best Local Similarity 99.3%; Pred. No. 2.9e-184; Indels 0; Gaps 0;
Matches 604; Conservative 0; Mismatches 4;

QY 2 CGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVQMLAQAAEE 61
Db 1 CGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVDAEGHMTLRLRLKGVQMLAQAAEE 60
QY 62 HPLHGGTGAHTRWATHGEPSEANAPHVSEHIVVHNGIIEHPELREELKARGYTFVS 121
Db 61 HPLHGGTGAHTRWATHGEPSEANAPHVSEHIVVHNGIIEHPELREELKARGYTFVS 120
QY 122 ETDTEVIHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSLVI 181
Db 121 ETDTEVIHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSLVI 180
QY 182 GLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQY 241
Db 181 GLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVHIQILAC 301
Db 241 DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLELLSKVHIQILAC 300
QY 302 GTSYNSGMVSKYWPESLAGIPCDVEIASERFYSKSAVRNSLMTLSQSGETADTLAQLR 361
Db 301 GTSYNSGMVSKYWPESLAGIPCDVEIASERFYSKSAVRNSLMTLSQSGETADTLAQLR 360
QY 362 LSKELGYGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLVLLMLVAKLS 421
Db 361 LSKELGYGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSRGQDYPAL 481
Db 421 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSRGQDYPAL 480
QY 482 EGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 541
Db 481 EGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPITFTVPLQLLAYHVALIKGTDVDDQPRN 601

```

```

Db      541 GQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDVQPRN 600
QY      602 LAKSVTVE 609
      |||||
Db      601 LAKSVTVE 608

RESULT 7
Q3YN3 SHISS
ID      Q3YN3_SHISS PRELIMINARY; PRT; 609 AA.
AC      Q3YN3;
DT      27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT      27-SEP-2005, sequence version 1.
DT      07-FEB-2006, entry version 5.
DE      L-glutamine:D-fructose-6-phosphate aminotransferase.
GN      Name=glms; OrderedLocNames=SSO_3890; ORFNames=SSO_3890;
OS      Shigella sonnei (strain Ss046).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Shigella.
OX      NCBI_TaxID=300269;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      PubMed=16275786; DOI=10.1093/nar/gki954;
RA      Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA      Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA      Nie H., Peng J., Xu J., Wang J., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA      Qiang B., Hou Y., Yu J., Jin Q.;
RT      "Genome dynamics and diversity of Shigella species, the etiologic
RT      agents of bacillary dysentery.";
RL      Nucleic Acids Res. 33:6445-6458(2005).
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DB      EMBL; CP000038; AAZ90429.1; -; Genomic DNA.
DR      GO; GO:0005737; C:cytoplasm; IEA.
DR      GO; GO:0004360; P:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR      GO; GO:0005529; P:sugar binding; IEA.
DR      GO; GO:0016740; P:transferase activity; IEA.
DR      GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR000593; GATase_2.
DR      InterPro; IPR005855; Glms_trans.
DR      InterPro; IPR001347; SIS.
DR      Pfam; PF00310; GATase 2; 1.
DR      Pfam; PF01380; SIS; 2.
DR      TIGRFAMs; TIGR01135; glms; 1.
DR      PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN 1.
KW      Aminotransferase; Complete proteome; Transferase.
SQ      SEQUENCE 609 AA; 66844 MW; 9BE75125C4D348D4 CRC64;

Query Match      98.9%; Score 3055; DB 2; Length 609;
Best Local Similarity 99.2%; Pred. No. 5, 1e-184;
Matches 604; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 MCGIVGAIQAQRDAVAILLEGLRRLRYGYDSAGLAVVDTEGHMTLRLRLGKQVQLAAAE 60
      |||||
Db      1 MCGIVGAIQAQRDAVAILLEGLRRLRYGYDSAGLAVDAEGHMTLRLRLGKQVQLAAAE 60
      |||||
QY      61 EHPHGGTGIAHTRWATGPESEVNAHPVSHVVVHNGIIEHPELREELKARGYTFV 120
      |||||
Db      61 EHPHGGTGIAHTRWATGPESEVNAHPVSHVVVHNGIIEHPELREELKARGYTFV 120
      |||||
QY      121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMIDSRHPTDTLAARSGSLV 180
      |||||
Db      121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMIDSRHPTDTLAARSGSLV 180
      |||||
QY      181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDTAEITRRSVNIFDKTGAEVKQRQDIESNLQ 240
      |||||
Db      181 IGLGNGENFIASDQALLPVTRRRFIFLEEGDTAEITRRSVNIFDKTGAEVKQRQDIESNLQ 240
      |||||
QY      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300
      |||||

```

```

Db      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300
      |||||
QY      301 CGTSYNSGMVSRYPESLAGIPCDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL 360
      |||||
Db      301 CGTSYNSGMVSRYPESLAGIPCDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL 360
      |||||
QY      361 RLSKELGYLGLSLAICNVFGSSILVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
      |||||
Db      361 RLSKELGYLGLSLAICNVFGSSILVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
      |||||
QY      421 SRLKGLDASIEHDIIVHGIQALPSRIEOMLSQDKRIEALAEADPSDKHHLFLSRGQOYPTA 480
      |||||
Db      421 SRLKGLDASIEHDIIVHGIQALPSRIEOMLSQDKRIEALAEADPSDKHHLFLSRGQOYPTA 480
      |||||
QY      481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVTVVAPNNLELLEKLSNIEVRRAR 540
      |||||
Db      481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVTVVAPNNLELLEKLSNIEVRRAR 540
      |||||
QY      541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDVQPR 600
      |||||
Db      541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDVQPR 600
      |||||
QY      601 NLAKSVTVE 609
      |||||
Db      601 NLAKSVTVE 609
      |||||

RESULT 8
GLMS_ECOL6
ID      GLMS_ECOL6 STANDARD; PRT; 608 AA.
AC      Q8FBT4;
DT      06-JUN-2003, integrated into UniProtKB/Swiss-Prot.
DT      06-JUN-2003, sequence version 2.
DT      07-MAR-2006, entry version 17.
DE      Glucosamine-fructose-6-phosphate aminotransferase [isomerizing]
DE      (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE      phosphate amidotransferase) (GfAT) (L-glutamine-D-fructose-6-phosphate
DE      amidotransferase) (Glucosamine-6-phosphate synthase).
GN      Name=glms; OrderedLocNames=c4654;
OS      Escherichia coli O6.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=217992;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX      MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA      Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA      Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA      Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA      Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT      "Extensive mosaic structure revealed by the complete genome sequence
RT      of uropathogenic Escherichia coli.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC      !- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC      converting fructose-6p into glucosamine-6p using glutamine as a
CC      nitrogen source (By similarity).
CC      !- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC      glutamate + D-glucosamine 6-phosphate.
CC      !- SUBUNIT: Homodimer (By similarity).
CC      !- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC      !- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC      !- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DB      EMBL; AE016769; AAN83086.1; -; Genomic DNA.
DR      HSSP; P17169; 1MOS.
DR      SMR; Q8FBT4; 1-239, 243-608.
DR      GenomeReviews; AE014075_GR; c4654.

```



```
DR BioCyc; ECOL199310:C4654-MONOMER; -
DR HAMAP; MF_00164; -; 1
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR005855; GIMS_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; GATase_2; 1.
DR TIGRFAMs; TIGR01135; gims; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FT amidotransferase [isomerizing].
FT FT /FtId=PRO_0000135323.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66647 MW; 44DADB5072C65D7 CRC64;

Query Match 98.7%; Score 3049; DB 1; Length 608;
Best Local Similarity 99.0%; Pred. No. 1.2e-183;
Matches 602; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRLGKQVQLAAQAAEE 61
DB 1 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRLGKQVQLAAQAAEE 60
QY 62 HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFVS 121
DB 61 HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFVS 120
QY 122 ETTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSLVI 181
DB 121 ETTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSLVI 180
QY 182 GLGWFNFASDQALAPVTRRTFFLEEGDIAETTRSVNIFDKTGAEVKRDIESNLQY 241
DB 181 GLGWFNFASDQALAPVTRRTFFLEEGDIAETTRSVNIFDKTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYHYMQEIIYEQPNAINKTTLTGRIHQGVQDLSBELGPNADLLSKVHEHIQILAC 301
DB 241 DAGDKGIYHYMQEIIYEQPNAINKTTLTGRIHQGVQDLSBELGPNADLLSKVHEHIQILAC 300
QY 302 GTSYNSGWSRYWFESLAGIPCDVEIASFEFRYKSAVRNLSMLTTLQSGETADTLAQLR 361
DB 301 GTSYNSGWSRYWFESLAGIPCDVEIASFEFRYKSAVRNLSMLTTLQSGETADTLAQLR 360
QY 362 LSKELGVLGSLAICNVPGSSIVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 421
DB 361 LSKELGVLGSLAICNVPGSSIVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIHVHGLQALPSRIEQMLSDQDKRIEALAEFSDKHAFSLSRGQDQYPIAL 481
DB 421 RLKGLDASIEHDIHVHGLQALPSRIEQMLSDQDKRIEALAEFSDKHAFSLSRGQDQYPIAL 480
QY 482 EGALKKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELKELKSNIEEVVARG 541
DB 481 EGALKKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELKELKSNIEEVVARG 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTDVDDPRN 601
DB 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFTYVPLQLLAYHVALIKGTDVDDPRN 600
QY 602 LAKSVTVE 609
DB 601 LAKSVTVE 608

RESULT 9
GLMS SALPA STANDARD; PRT; 608 AA.
ID GLMS SALPA
AC QSPK79;
```

```
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 12.
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Names:glms; OrderedLocNames=SPA3700;
OS Salmonella paratyphi-a
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=ATCC 9150 / SARB42;
RX PubMed=1531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du P., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P., Florea L.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RL Nat. Genet. 36:1268-1274(2004).
CC -1- FUNCTION: catalyzes the first step in hexosamine metabolism, as a
CC . converting fructose-6p into glucosamine-6p using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFAT subfamily.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; CP000026; AAV79492.1; -; Genomic_DNA.
CC SMR; QSPKV9; 1-239, 243-608.
CC GenomeReviews; CP000026_GR; SPA3700.
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; GIMS_trans.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC Pfam; PF01380; SIS; 2.
CC TIGRFAMs; TIGR01135; gims; 1.
CC PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FT amidotransferase [isomerizing].
FT FT /FtId=PRO_0000135373.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66719 MW; EE20C7F1996B742B CRC64;

Query Match 98.4%; Score 3040; DB 1; Length 608;
Best Local Similarity 98.7%; Pred. No. 4.5e-183;
Matches 600; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRLGKQVQLAAQAAEE 61
DB 1 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRLGKQVQLAAQAAEE 60
QY 62 HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFVS 121
DB 61 HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFVS 120
```

```
QY 122 ETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLVI 181
Db 121 ETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLVI 180
QY 182 GLGCMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIDKTAEGAEVKRQDIESNQY 241
Db 181 GLGCMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIDKTAEGAEVKRQDIESNQY 240
QY 242 DAGDKGIYCHYMOKIEYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 301
Db 241 DAGDKGIYCHYMOKIEYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 300
QY 302 GTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAIGLR 361
Db 301 GTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAIGLR 360
QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLMTWAGTEIGVASTKATFTTLLMLVAKLS 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLMTWAGTEIGVASTKATFTTLLMLVAKLS 420
QY 422 RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRGDQYPIAL 481
Db 421 RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRGDQYPIAL 480
QY 482 EGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 541
Db 481 EGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 542 GQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPN 601
Db 541 GQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPN 600
QY 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 10
ID Q57HY2_SALCH PRELIMINARY; PRT; 609 AA.
AC Q57HY2;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=gIms; OrderedLocusNames=SC3774; ORFNames=SCH_3774;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AE017220; RAX67680.1; -; Genomic_DNA.
DR SMR; Q57HY2; 2-240, 244-609.
DR GO; GO:0005737; C:cycloplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000583; GATase_2.
```

```
DR InterPro; IPR005855; GImS_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; gImS; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN 1.
KW Aminotransferase; Complete proteome; Transferrase.
SQ SEQUENCE 609 AA; 66849 NW; EIDIFF6FIAF35F5C CRC64;

Query Match          98.4%; Score 3040; DB 2; Length 609;
Best Local Similarity 98.4%; Pred. No. 4.6e-183;
Matches 599; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAORDVABAILLEGRLRLLEYRGYDSAGLAVVDTGHTMLRLRLGKQVQLAAQAE 60
Db 1 MCGIVGAIQAORDVABAILLEGRLRLLEYRGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAQAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEENHEPLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTAEGAEVKRQDIESNQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTAEGAEVKRQDIESNQ 240
QY 241 YDAGDKGIYCHYMOKIEYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 300
Db 241 YDAGDKGIYCHYMOKIEYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAIGLR 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAIGLR 360
QY 361 RLSKELGYLGLSLAICNVPGSSLVRESDLMTWAGTEIGVASTKATFTTLLMLVAKL 420
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLMTWAGTEIGVASTKATFTTLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRGDQYPIA 480
Db 421 ARKGLDTSIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
GLMS_SALTY
ID GLMS_SALTY STANDARD; PRT; 608 AA.
AC Q8ZKX1;
DT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT 11-JUL-2002, sequence version 2.
DT 07-MAR-2006, entry version 21.
DE Glucosamine-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPA) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=gImS; OrderedLocusNames=STM3861;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
```


FT	REGION	1	241	Glutamine amidotransferase.
FT	ACT_SITE	1	1	CATase (By similarity).
FT	ACT_SITE	604	604	Isomerization Fru-6P (By similarity).
SQ	SEQUENCE	609 AA; 6686 MW; BFE71B8E5DFD23B5D CRC64;		
	Query Match	86.4%; Score 2667.5; DB 1; Length 609;		
	Best Local Similarity	84.7%; Pred. No. 1.4e-159;		
	Matches	516; Conservative 46; Mismatches 46; Indels 1; Gaps 1;		
Qy	2	CGIVGAIQADVAIIELEGRLRLRYGYDSAGLAVDTEGHMTLRRLRGKQVLAQAAEE 61		
Db	1	CGIVGAVAQADVAIIELEGRLRLRYGYDSAGLAVDSEGHVARLRRLRGKQVLSQAAEE 60		
Qy	62	HLPLGGTGIAHTRWATGCESEVNAHPVSEHIVVNGIIEHNEPEELKARGITFVS 121		
Db	61	HELHGGTGIAHTRWATGCESEENAPHISEHITVHNGIIEHNEPEELMIGRGYFVS 120		
Qy	122	ETDTEVIAHLVNWELKO-CGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180		
Db	121	ETTEVVAHLVHFQKQNGTIVVVKVIFQLRGAGVWVLDNRDSSVLVAARSGSPLV 180		
Qy	181	IGLGMGNFIASDQALLPVTRPFIIEEGDIAEITRRSVNI FDKTGAEVKRDIESNLQ 240		
Db	181	IGRGVGNFIASDQALLPVTRRPMFLEEGDVAEITRRDVRVFDKSGQLATREEIESKVS 240		
Qy	241	YDAGDKGIYCHYQKEIYEQNAIKNTLTGRI SHGOVDLSLGNADNELLKVEHIIOLA 300		
Db	241	YDAGDKGAYRHQYQKEIYEQPMATKNTLEGRFSHGELNLSLGNADNELLKVEHVQIIA 300		
Qy	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRKRSVNRNLSMITLSQSGETADTLA 360		
Db	301	CGTSYNSGMVSRVWFELAGIPCDVEIASFPRKRPVNRNLSMITLSQSGETADTLAAL 360		
Qy	361	RLSKELGYLSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFQTQTTLMLVLAK 420		
Db	361	RLSKELGYLSGLAICNVAGSSLVRESDMALMTRAGVEIGVASTKAFQTQTTLMLVARV 420		
Qy	421	SLRKLGDASTEHDIHVHGLQALPSRIEQLMSQDKRIEALAEDFSXKHALFLSRGDQYPIA 480		
Db	421	GRUGMDAQTEHDIHVHGLQALPARIEQLMSQDKLIESLAEFGFSXKHALFLGRGDQYPIA 480		
Qy	481	LEGALKLKETSITHAEAYAGELKHGFLALIDAMPVIVVAPNNELEKLSNIEEVRAR 540		
Db	481	MEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNELEKLSNIEEVRAR 540		
Qy	541	GGOLYVFADQADGVSSDNNHIIEMPHVEVIAPIFYTVPLQLLXHVHAIKGTVDVQPR 600		
Db	541	GGELYVFADEADGTSSENKIIPLPHIEEVIAPIFYTVPLQLLSYHVALIKGTVDVQPR 600		
Qy	601	NLAKSVTIVE 609		
Db	601	NLAKSVTIVE 609		
	RESULT 14			
GLMS_YERPE	STANDARD;	PRT;	608 AA.	
AC	Q8Z9S8;			
DT	11-JUL-2002,	integrated into UniProtKB/Swiss-Prot.		
DT	11-JUL-2002,	sequence version 2.		
DT	07-MAR-2006,	entry version 32.		
DE	Glucosamine-6-phosphate	aminotransferase [isomerizing]		
DE	(EC 2.6.1.16)	(hexosephosphate aminotransferase) (D-fructose-6-phosphate amidotransferase) (GPAT) (L-glutamine-D-fructose-6-phosphate		
DE	Name=glms;	OrderedLocusNames=YPO4118, y4132, YP4025;		
OS	Yersinia pestis			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Yersinia.			
OX	NCBI_TaxID=632;			
RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RP	STRAIN=CO-92 / Biovar Orientalis;			

RX	MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;			
RA	Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,			
RA	Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,			
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,			
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,			
RA	Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,			
RA	Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,			
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;			
RT	"Genome sequence of Yersinia pestis, the causative agent of plague."			
RL	Nature 413:523-527(2001).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=KIMS / Biovar Mediaevalis;			
RX	MEDLINE=22137863; PubMed=12142430;			
RX	DOI=10.1128/JB.184.16.4601-4611.2002;			
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,			
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,			
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,			
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,			
RA	Perry R.D.;			
RT	"Genome sequence of Yersinia pestis KIM."			
RL	J. Bacteriol. 184:4601-4611(2002).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=91001 / Biovar Mediaevalis;			
RX	PubMed=15368893; DOI=10.1093/dnares/11.3.179;			
RA	Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,			
RA	Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,			
RA	Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,			
RA	Yang H., Wang J., Huang P., Yang R.;			
RT	"Complete genome sequence of Yersinia pestis strain 91001, an isolate			
RT	avirulent to humans."			
RL	DNA Res. 11:179-197(2004).			
CC	-I- FUNCTION: Catalyzes the first step in hexosamine metabolism,			
CC	converting fructose-6p into glucosamine-6p using glutamine as a			
CC	nitrogen source (By similarity).			
CC	-I- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-			
CC	glutamate + D-glucosamine 6-phosphate.			
CC	-I- SUBCELLULAR LOCATION: Cytoplasm (By similarity).			
CC	-I- SIMILARITY: In the C-terminal section; belongs to the SIS family.			
CC	GPAT subfamily.			
CC	-I- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.			
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms			
CC	Distributed under the Creative Commons Attribution-NoDerivs License			
CC	-----			
EMBL	AJ414160; CAC93567.1; -; Genomic DNA.			
EMBL	AE014014; AAM87674.1; -; Genomic DNA.			
EMBL	AE017142; AAS64164.1; -; Genomic DNA.			
PIR	AB0500; AB0500.			
HSSP	PI7169; IMOS.			
SNR	Q829S8; 1-239, 243-608.			
GenomeReviews	AE009952 GR; Y4132.			
GenomeReviews	AE017042 GR; YP4025.			
GenomeReviews	AL590842 GR; YPO4118.			
BioCyc	YPES187410:Y4132-MONOMER; -.			
BioCyc	YPES632:YPO4118-MONOMER; -.			
HMAP	MF 00164; -; 1.			
InterPro	IPR000583; GATase_2.			
InterPro	IPR005855; Glms_trans.			
InterPro	IPR001347; SIS_trans.			
Pfam	PF00310; GATase_2; 1.			
Pfam	PF01380; SIS; 2.			
TIGRFAMs	TIGR01135; glms; 1.			
PROSITE	PS00443; GATASE_TYPE II; 1.			
Aminotransferase	Complete proteome; Glutamine amidotransferase;			
Transferase.				
INIT_MET	0 0 By similarity.			
CHAIN	1 608 Glucosamine--fructose-6-phosphate			
	aminotransferase [isomerizing].			
	/FTid=PRO 0000135418.			
REGION	1 240 Glutamine amidotransferase.			
ACT_SITE	1 1 GATase (By similarity).			

```
FT ACT_SITE 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.2%; Score 2663; DB 1; Length 608;
Matches 515; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 2 CGIVGAIAQORDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRRLRRKGVOMLAQAABE 61
Db 1 CGIVGAVAQORDIAEILLEGRLRLEVRGYDSAGLAVVDSEGLHTRLRVRKGVHVALSDAAEK 60
QY 62 HPLHGGTGIATRWATHGEPSEVNAHPVSEHVVVHNGIIEHNPRLREELKARGYTFVS 121
Db 61 QDLHGGTGIATRWATHGEPSEANAHPHVSDYISVVHNGIIEHNPRLRELLISRGYRFS 120
QY 122 ETDTEVIAHLVNWELKQGTGLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPVLI 181
Db 121 ETDTEVIAHLVHWEQGGSLLEVVKRVIPQLRGAYGTVMDSRPSRLIAARSGSPVLI 180
QY 182 GLGMGENFIASDQALLPVTFRFIFLEEGDIAETRRSVNIPDKTGAEVKRDQIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTFRFIFLEEGDVVEVTRRSISIFDKQGNAIERPEIESOVQY 240
QY 242 DAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC 301
Db 241 DAGDKGIYRHYMQKEIYEQPNAIKNTLEGRLSHGMDLSELGPKADALLAEVQHIQIILAC 300
QY 302 GTSYNSGMVSRVYWFESLAGIPCDEIASEFRYKSAVRNSLMTLSQSGETADTLAAGR 361
Db 301 GTSYNSGMVSRVYWFESLAGVPCDEIASEFRYKSAVRNSLMTLSQSGETADTLAALR 360
QY 362 LSKELGYLGSILAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKUS 421
Db 361 LSKELGYLGSILAICNVAGSSLVRESDLALMTKAGTEIGVASTKFTTQTLTVLLMLVGRIG 420

RESULT 15
GLMS YERPS
ID GLMS YERPS STANDARD; PRT; 608 AA.
AC Q663R1;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DE Glucosamine-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (G6P) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=g6ms; OrderedLocustNames=YPTB3964;
OS Versinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
```

```
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebuch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Dethise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Versinia pestis through whole-genome
RT comparison with Versinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFAT subfamily.
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BX936398; CAH23202.1; -; Genomic_DNA.
DR SMR; Q663R1; 1-239, 243-608.
DR GenomeReviews; BX936398 GR; YPTB3964.
DR HAMAP; MF_00164; -; 1.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR005855; G6ms_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR TIGRFAMs; TIGR01135; g6ms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608
FT Glucosamine--fructose-6-phosphate
FT amidotransferase [isomerizing].
FT /FTID=PRO_0000135419.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.2%; Score 2663; DB 1; Length 608;
Matches 515; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 2 CGIVGAIAQORDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRRLRRKGVOMLAQAABE 61
Db 1 CGIVGAVAQORDIAEILLEGRLRLEVRGYDSAGLAVVDSEGLHTRLRVRKGVHVALSDAAEK 60
QY 62 HPLHGGTGIATRWATHGEPSEVNAHPVSEHVVVHNGIIEHNPRLREELKARGYTFVS 121
Db 61 QDLHGGTGIATRWATHGEPSEANAHPHVSDYISVVHNGIIEHNPRLRELLISRGYRFS 120
QY 122 ETDTEVIAHLVNWELKQGTGLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPVLI 181
Db 121 ETDTEVIAHLVHWEQGGSLLEVVKRVIPQLRGAYGTVMDSRPSRLIAARSGSPVLI 180
QY 182 GLGMGENFIASDQALLPVTFRFIFLEEGDIAETRRSVNIPDKTGAEVKRDQIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTFRFIFLEEGDVVEVTRRSISIFDKQGNAIERPEIESOVQY 240
QY 242 DAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC 301
Db 241 DAGDKGIYRHYMQKEIYEQPNAIKNTLEGRLSHGMDLSELGPKADALLAEVQHIQIILAC 300
QY 302 GTSYNSGMVSRVYWFESLAGIPCDEIASEFRYKSAVRNSLMTLSQSGETADTLAAGR 361
Db 301 GTSYNSGMVSRVYWFESLAGVPCDEIASEFRYKSAVRNSLMTLSQSGETADTLAALR 360
QY 362 LSKELGYLGSILAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKUS 421
Db 361 LSKELGYLGSILAICNVAGSSLVRESDLALMTKAGTEIGVASTKFTTQTLTVLLMLVGRIG 420
```

Qy	422	RLKGLDASIEHDI VHGI QALPSRIEOMLSODKRIEALAEDEFSKXHALFLSRGDVPIAL	481
Db	421	KLKGADASLEHDI VHALQALPARIEQMLSLDKTIEALAEDEFSKXHALFLSRGDVPIAM	480
Qy	482	EGALKLKEISYIHAEAYAAGELKHGPIALIDADMPVIVVAPNPNELLEKLKSNIEEVRARG	541
Db	481	EGALKLKEISYIHAEAYAAGELKHGPIALIDADMPVIVVAPNPNELLEKLKSNIEEVRARG	540
Qy	542	GQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFYTVPLOLLAYHVHVALIKGTDVDQPRN	601
Db	541	GLLYVFADQDAGFTDSEGMKIIQLPHVEEIIAPIFYTVPLOLLSYHVHVALIKGTDVDQPRN	600
Qy	602	LAKSVTVE	609
Db	601	LAKSVTVE	608

Search completed: June 14, 2006, 15:27:41
Job time : 308.293 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:21:37 ; Search time 41.2344 Seconds
(without alignments)
1421.048 Million cell updates/sec

Title: US-10-612-779-6
Perfect score: 3089
Sequence: 1 MCGIVGAIQAQDVAEILLEG.....LIKGTVDQPRNLAKSVTVE 609

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Pirl:*
2: Pirl:*
3: Pirl:*
4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3068	99.3	609	1 XNECGM	glutamine-fructose
2	3064	99.2	609	2 H86058	hypothetical prote
3	3057	99.0	609	2 G91212	hypothetical prote
4	3035	98.3	609	2 AB0955	glutamine-fructose
5	2668	86.4	609	2 AB0500	glutamine-fructose
6	2256.5	73.0	610	2 D64067	glutamine-fructose
7	2250.5	72.9	610	2 B82316	glutamine-fructose
8	1983	64.2	611	2 F82951	glutamine-fructose
9	1856	60.1	621	2 A84933	glutamine-fructose
10	1688	54.6	635	2 E71272	glutamine-fructose
11	1681	54.4	609	2 A82844	glutamine-fructose
12	1579	51.1	611	2 T45493	glutamine-fructose
13	1569.5	50.8	612	2 H81246	glutamine-fructose
14	1565.5	50.7	612	2 B81246	glutamine-fructose
15	1435	46.5	612	2 AD3595	glutamine-fructose
16	1417.5	45.9	608	2 B95322	nodM Glutamine ami
17	1413.5	45.8	608	2 C97575	glutamine-fructose
18	1413.5	45.8	608	2 AC2796	hypothetical prote
19	1404.5	45.5	604	2 S67993	glutamine-fructose
20	1362.5	44.1	606	2 F75536	glutamine-fructose
21	1350.5	43.7	592	2 D70327	glutamine-fructose
22	1333.5	43.2	608	1 S01040	glutamine-fructose
23	1323.5	42.8	606	2 B87263	hypothetical prote
24	1321	42.8	605	2 S16561	glutamine-fructose
25	1320.5	42.7	598	2 F75212	glutamine-fructose
26	1303.5	42.2	606	2 B72412	glutamine-fructose
27	1287.5	41.7	608	2 C96919	glutamine-fructose
28	1284	41.6	615	2 T35569	glutamine-fructose
29	1282.5	41.5	601	2 D71248	glutamine-fructose

ALIGNMENTS

RESULT 1

XNECGM

glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Escherichia coli
N; Alternate names: glucosamine-6-phosphate synthase; glucosaminophosphate isomerase (glu
C; Species: Escherichia coli
C; Date: 31-Mar-1990 #sequence_revision 17-Oct-1997 #text_change 09-Jul-2004
R; Accession: B65176; A30389; I41219; S17839; Q90513
R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID:97426617; PMID:9278503
A; Accession: B65176
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-609 <BLAT>
A; Cross-references: UNIPROT:P17169; UNIPARC:UPI0000153CA3; GB:AE000450; GB:U00096; NID:G
A; Experimental source: strain K-12, substrain MG1655
R; Walker, J.E.; Gay, N.J.; Saraste, M.; Eberle, A.N.
Biochem. J. 224, 799-815, 1984
A; Title: DNA sequence around the Escherichia coli unc operon. Completion of the sequence
A; Reference number: A30389; MUID:85121806; PMID:6395859
A; Accession: A30389
A; Molecule type: DNA
A; Residues: 1-418, 'NV', 421-609 <WAL>
A; Cross-references: UNIPARC:UPI000016F5F4; GB:X01631; NID:G43256; PIDN:CAA25785.1; PID:G
R; McKown, R.L.; Orle, K.A.; Chen, T.; Craig, N.L.
J. Bacteriol. 170, 352-358, 1988
A; Title: Sequence requirements of Escherichia coli attTn7, a specific site of transposon
A; Reference number: I41219; MUID:88086894; PMID:2826397
A; Accession: I41219
A; Status: translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 597-609 <RES>
A; Cross-references: UNIPARC:UPI000016F18D; GB:M19880; NID:G146074; PIDN:AAA23836.1; PID:G
R; Golinielli-Pimpaneau, B.; Badet, B.
Eur. J. Biochem. 201, 175-182, 1991
A; Title: Possible involvement of Lys603 from Escherichia coli glucosamine-6-phosphate syr
A; Reference number: S17839; MUID:92007872; PMID:1915361
A; Accession: S17839
A; Molecule type: protein
A; Residues: 49-50, 'X', 52-53; 218, 219, 220-223, 'X', 225-231; 489-493; 504-508, 'K', 'K', 601-609
A; Cross-references: UNIPARC:UPI000011EBBE; UNIPARC:UPI000011ECBE; UNIPARC:UPI000011ECBE;
4A7
C; Genetics:
A; Gene: glms
A; Map position: 4 min
C; Function:
A; Description: catalyzes the formation of D-glucosamine-6-phosphate from the amido group
A; Pathway: glucosamine biosynthesis
A; Note: Glucosamine-6-phosphate is used in the biosynthesis of amino sugars of asparagin

glutamine-fructose
glutamine-fructose
glutamine-fructose
probable glutamine
hypothetical prote
glutamine-fructose
L-glutamine-D-fruc
glutamine-fructose
glutamine-fructose
glutamine-fructose
glutamine-fructose
hypothetical prote
glutamine-fructose
glutamine-fructose
L-glutamine-D-fruc
glutamine-fructose

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-609/Product: Glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F:2/Active site: Cys #status predicted

Query Match 99.3%; Score 3068; DB 1; Length 609;
Best Local Similarity 99.5%; Pred. No. 6e-191;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRLGKQVQLAAQAAE 60
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRLGKQVQLAAQAAE 60

QY 61 EPHLHGCGTGAIAHTRWATHGEPSEVNAHPHVSHIIVVHNGIIEENHEPLREELKARGYTFV 120
DB 61 EPHLHGCGTGAIAHTRWATHGEPSEVNAHPHVSHIIVVHNGIIEENHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLNVNWKQGGTLREAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180
DB 121 SETDTEVIAHLNVNWKQGGTLREAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180

QY 181 IGLGNGENFIASDQALLPVTTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
DB 181 IGLGNGENFIASDQALLPVTTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEPRYKSAVRNLSLMTTLSQSGTADTLAGL 360
DB 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEPRYKSAVRNLSLMTTLSQSGTADTLAGL 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIYVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFFLSRGQDQYPTA 480
DB 421 SRLKGLDASIEHDIYVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFFLSRGQDQYPTA 480

QY 481 LEGALKLKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNLEKLNIEEVRAR 540
DB 481 LEGALKLKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNLEKLNIEEVRAR 540

QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDQPR 600

QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 2
H86058
hypothetical protein glms [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H86058
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaener, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H86058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <STO>
A:Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI0000165991; GB:AE005174; NID:g12518583; F
C:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: glms

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
Query Match 99.2%; Score 3064; DB 2; Length 609;
Best Local Similarity 99.3%; Pred. No. 1.1e-190;
Matches 605; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRLGKQVQLAAQAAE 60
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRLGKQVQLAAQAAE 60

QY 61 EPHLHGCGTGAIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
DB 61 EPHLHGCGTGAIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLNVNWKQGGTLREAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180
DB 121 SETDTEVIAHLNVNWKQGGTLREAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180

QY 181 IGLGNGENFIASDQALLPVTTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
DB 181 IGLGNGENFIASDQALLPVTTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEPRYKSAVRNLSLMTTLSQSGTADTLAGL 360
DB 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEPRYKSAVRNLSLMTTLSQSGTADTLAGL 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIYVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFFLSRGQDQYPTA 480
DB 421 SRLKGLDASIEHDIYVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFFLSRGQDQYPTA 480

QY 481 LEGALKLKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNLEKLNIEEVRAR 540
DB 481 LEGALKLKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNLEKLNIEEVRAR 540

QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDQPR 600

QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 3
G91212
hypothetical protein ECs4671 [imported] - Escherichia coli (strain O157:H7, substrain RIN
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: G91212
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G91212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <HAY>
A:Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI000016552F; GB:BA000007; PIDN:BAE38094.1;
C:Experimental source: strain O157:H7, substrain RIND 0509952
C:Genetics:
A:Gene: ECs4671
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
Query Match 99.0%; Score 3057; DB 2; Length 609;

Best Local Similarity 99.2%; Pred. No. 3.1e-190;		Matches 604; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
Qy	1	MC GIVGAI AQRDVAEILLEGRLRELYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE	60
Db	1	MC GIVGAI AQRDVAEILLEGRLRELYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE	60
Qy	61	EHP LHGGTGTATHRWATHGEPSEVNAHPHVSEHVVVHNGI IENHEPRLBELKARGYTFV	120
Db	61	EHP LHGGTGTATHRWATHGEPSEVNAHPHVSEHVVVHNGI IENHEPRLBELKARGYTFV	120
Qy	121	SETDTEVIAHLVNWELKQGGLTREA VLR AIPQLRGAYGTVMDSRHPD TLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNWELKQGGLTREA VLR AIPQLRGAYGTVMDSRHPD TLLAARSGSPLV	180
Qy	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ	240
Db	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ	240
Qy	241	YDAGDKGIYCHYWKQEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYWKQEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLLSKVEHIQILA	300
Qy	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFVRKSAVRNSLMITLSQSGETADTLAGL	360
Db	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFVRKSAVRNSLMITLSQSGETADTLAGL	360
Qy	361	RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKAF TQTLTVLLMLVAKL	420
Db	361	RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKAF TQTLTVLLMLVAKL	420
Qy	421	SR LKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAE DFDKXHALFLSRGDQYPIA	480
Db	421	SR LKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAE DFDKXHALFLSRGDQYPIA	480
Qy	481	LEGALKLKEISYTHAEAYAGELKHG PLALIDAMPVIVVAPNNLEKLSNIEEVRAR	540
Db	481	LEGALKLKEISYTHAEAYAGELKHG PLALIDAMPVIVVAPNNLEKLSNIEEVRAR	540
Qy	541	GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDVDQPR	600
Db	541	GGQLYVPAEQDAGFVSSDNNHIIEMPNVEEVIAPFYTVPLQLLHYHVALIKGTDVDQPR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	

RESULT 4
AB0955
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Salmonella enteritidis
N;Alternate names: Glucosamine-fructose-6-P aminotransferase (mismomer)
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 28-Jul-2003
C;Accession: AB0955
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB0955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-609 <PAR>
A;Cross-references: UNIPARC:UPI000005A683; GB:AL513382; PIDN:CAD03134.1; PID:G16504770;
C;Genetics:
A;Gene: STY3917
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase

Query Match 98.13%; Score 3035; DB 2; Length 609;		Best Local Similarity 98.0%; Pred. No. 8.2e-189;	
Matches 597; Conservative 5; Mismatches 7; Indels 0; Gaps 0;			
Qy	1	MC GIVGAI AQRDVAEILLEGRLRELYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE	60
Db	1	MC GIVGAI AQRDVAEILLEGRLRELYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE	60
Qy	61	EHP LHGGTGTATHRWATHGEPSEVNAHPHVSEHVVVHNGI IENHEPRLBELKARGYTFV	120
Db	61	EHP LHGGTGTATHRWATHGEPSEVNAHPHVSEHVVVHNGI IENHEPRLBELKARGYTFV	120
Qy	121	SETDTEVIAHLVNWELKQGGLTREA VLR AIPQLRGAYGTVMDSRHPD TLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNWELKQGGLTREA VLR AIPQLRGAYGTVMDSRHPD TLLAARSGSPLV	180
Qy	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ	240
Db	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ	240
Qy	241	YDAGDKGIYCHYWKQEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYWKQEIYEOPNAIKNTLTGRI SHGQVDLSLGNADLLSKVEHIQILA	300
Qy	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFVRKSAVRNSLMITLSQSGETADTLAGL	360
Db	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFVRKSAVRNSLMITLSQSGETADTLAGL	360
Qy	361	RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKAF TQTLTVLLMLVAKL	420
Db	361	RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKAF TQTLTVLLMLVAKL	420
Qy	421	SR LKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAE DFDKXHALFLSRGDQYPIA	480
Db	421	SR LKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAE DFDKXHALFLSRGDQYPIA	480
Qy	481	LEGALKLKEISYTHAEAYAGELKHG PLALIDAMPVIVVAPNNLEKLSNIEEVRAR	540
Db	481	LEGALKLKEISYTHAEAYAGELKHG PLALIDAMPVIVVAPNNLEKLSNIEEVRAR	540
Qy	541	GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDVDQPR	600
Db	541	GGQLYVFDQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDVDQPR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	

RESULT 5
AB0500
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB0500
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, N.; Moule, S.; O'Gaora, P.
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AB0500
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-609 <KUR>
A;Cross-references: UNIPROT:Q8Z9S8; UNIPARC:UPI0000165B7A; GB:AL590842; PIDN:CAC93567.1;
C;Genetics:
A;Gene: glms
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase

Query Match 86.4%; Score 2668; DB 2; Length 609;
Best Local Similarity 84.7%; Pred. No. 4.9e-165;

Matches 516; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHTRLRRLGKVOMLAQAAE 60
 DB 1 MCGIVGAVQORDIAEILLEGRLRLEYRGYDSAGLAVVDSSEGLTRLRRLGKVHLSDAE 60

QY 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHSEYVHVHNGIENHENHLEPLREELKARGYTFV 120
 DB 61 KODLHGCTGIAHTRWATHGEPSEANAPHVSDYISVVHNGIENHENHLEPLREELLISRGYRFS 120

QY 121 SETDTEVIAHLVNNELKQGTTLREAVLRATPOLRCAGYGTVMDSRHPDPTLLAARSGSPLV 180
 DB 121 SETDTEVIAHLVHVEQQGSLLEVVKRVIPOLRCAGYGTVMDSRDPSSLIAARSGSPLV 180

QY 181 IGLWGENFTASDQALLPVTRRFIFLEEGDIAETIRRSVNIPOKTCAEVKRODIESNLQ 240
 DB 181 ICGVGENFTASDQALLPVTRRFIFLEEGDVVEVTRRSISIFDKQGNWATERFEISQVQ 240

QY 241 YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGQVDLSSELGPNADLLSKVEHIQILA 300
 DB 241 YDAGDKGIYRHYMOKEIYEQPMKNTLEGRLSHGMDLSSELGPKADALLAEVQHIQIIA 300

QY 301 CGTSYNGMSVRYWFESLAGIPCDVEIASPFRYKSAVRNSLMITLSQSGETADTLAQL 360
 DB 301 CGTSYNGMSVRYWFESLAGVPCDVEIASPFRYKSAVRNSLLITLSQSGETADTLAAL 360

QY 361 RLSKELGYLGSALICNVPGSSLVRESDLALMTNAGTIGVASTKFTTQTLTVLLMLVAKL 420
 DB 361 RLSKELGYLGSALICNVAGSSLVRESDLALMTKAGTIGVASTKFTTQTLTVLLMLVGR 420

QY 421 SRLKGLDASIEHDTIVHGLQALPSRIEQWLSQDKRIEALAEFDSKXHALFLSRGQDQYPIA 480
 DB 421 GKLGXGASLEHDTVHVALQALPARIEQWLSLDKTIKIEALAEFDSKXHALFLSRGQDQYPIA 480

QY 481 LEGALKUKETSYIIHAEYAAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
 DB 481 MEGALKUKETSYIIHAEYAAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540

QY 541 GGQLYVFPADQDAGFVSSDNMHIIEMPHVEEVIAPFIVTLPQLLAYHVALIKGTVDVQDPR 600
 DB 541 GGLLYVFPADQDAGTDSGMKIIQUPHVEEIIAPFIVTLPQLLYSHVALIKGTVDVQDPR 600

QY 601 NLAKSVTIVE 609
 DB 601 NLAKSVTIVE 609

RESULT 6
 D64067
 glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Haemophilus influenzae
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
 C:Accession: D64067
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J. D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: D64067
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-610 <TIGR>
 A:Cross-references: UNIPROT:P44708; UNIPARC:UPI00000512BF; GB:U32726; GB:I42023; NID:g1918
 A:Note: named as homolog to a protein from Escherichia coli
 C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
 C:Keywords: aminotransferase; isomerase
 F:2-610/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
 F:2/Active site: Cys #status predicted

Query Match 73.0%; Score 2256.5; DB 2; Length 610;
 Best Local Similarity 72.0%; Pred. No. 2.2e-138;

Matches	439;	Conservative	73;	Mismatches	97;	Indels	1;	Gaps	1;
Qy	1	MCGIVGAIAORDVAEII	LEGRLRLEYRGYDSAGLAVVDTEGHMTRLRRLKGVQVLAQAAE	60					
Dd	1	MCGIVGAVAQRDVAEII	LINGLHRLREYRGYDSAGVAVINKQELQRIICLGKVKKALDEAVS	60					
Qy	61	EHLPHGGTGIATHTWATHGEPS	VNAHPVSHI VVVHNGI IENHEPLREELKARGYTFFV	120					
Dd	61	EKLPIGGTGIAHTRWATHGEPS	ETNAHPSSGTFAVVHNGI IENHEELRELLKSRYGVFL	120					
Qy	121	SETDTVEIAHLVNWELKOGGT	LREAVLRAIPOLRGAYGTVIMDSRRPDTLLAARSGSPLV	180					
Dd	121	SQDTVEIAHLVEMERTTD	SLLDRVKAKQLTGAYGMVMDSRPEHLVAARSGSPLV	180					
Qy	181	IGLGMGNFIASDQALLPVTR	RFFLEEGDIAETRISVNIFDKTGAEVKRODIESNLQ	240					
Dd	181	IGLGIGENFLASDQALLSV	TRRFIFLEEGDIAETRTRVDIYDTHGNKAKREIHESNLE	240					
Qy	241	YGADGKIYCHYMOKEIYEOP	NALKNTLTGRISHGQVDLSGLGNDADELLSKVEHIOLA	300					
Dd	241	NDAAEKGFRHFMOKEIYEOP	TALINTMEGRINHENVIVDSIGNGAKGLEKVEHIQIVA	300					
Qy	301	CGTSYNSGMTSYRVFESLAG	IPCDEVIESEPRYKSAVERNSLMITLSQSGETADTLAAGL	360					
Dd	301	CGTSYNAGWARVWFESLAG	VSCDVEIASFEPRYKFTVRPNSLIITLSQSGETADTLAAL	360					
Qy	361	RLSKELGYLGS LAI CNVP	GSLLRESDLALMTNACTEIGVASTKAFTQTUVTLLMLVAKL	420					
Dd	361	RLAKEGYMAALTICNVAGS	SLVRESDLAFMTRAGVEVGVAFTKTQLAALLMLVTAL	420					
Qy	421	SRLKG-IDASTEHDIVHGL	OALPSRIEQMLSODKRIEALAEDFSKXHALFLSRGDQYPI	479					
Dd	421	GKVKGHSIVSVEREBII	KAMQSLPAIEKALAPDTEIALAEDFAEKHHALFLGRGAFYPI	480					
Qy	480	ALEGALKLKEISYTHAEYA	AGELKHGPLALIDAMPVIWVAPNNELLEKLKSNIEEVRA	539					
Dd	481	AVERASKLKEISYTHAEYA	AGELKHGPLALIDAMPVIWVAPNNELLEKVKSNIEEVRA	540					
Qy	540	RGGOLYPADQADGFVSDD	NHIIEMPHVVEVIAPIFYTPVLQLLAHVALIKGTVDVDP	599					
Dd	541	RGGOLYPADKEAGFTPEG	MKIITMPKVNDIVAPIFYTPMQLLSYVALIKGTVDVDP	600					
Qy	600	RNLAKSVTVTE	609						
Dd	601	RNLAKSVTVTE	610						
RESULT	7								
E82316		glutamine-fructose-6-phosphate transaminase (isomerizing)	(EC 2.6.1.16) [similarity]	- v-					
C:	Species:	Vibrio cholerae							
C:	Date:	18-Aug-2000 #sequence_revision	20-Aug-2000 #text_change	09-Jul-2004					
C:	Accession:	E82316							
R:	Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;								
R:	Hardison, D.; Frolava, M.D.; Vamathevan, J.; Bacs, S.; Qin, H.; Dragoi, I.; Sellers, P.								
l,	R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.								
Nature	406,	477-483,	2000						
A:	Title:	DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.							
A:	Reference number:	A82035; PMID:20406833; PMID:10952301							
A:	Accession:	E82316							
A:	Status:	preliminary							
A:	Molecule type:	DNA							
A:	Residues:	1-610 <HEI>							
A:	Cross-references:	UNIPROT:O9KUM8; UNIPARC:UPI0000164B53; GB:AE004135; GB:AE003852; NID							
A:	Experimental source:	serogroup O1; strain N16961; biotype El Tor							
C:	Genetics:								
A:	Gene:	VC0487							
A:	Map position:	1							
C:	Superfamily:	glutamine-fructose-6-phosphate aminotransferase (isomerizing)							
C:	Keywords:	aminotransferase; isomerase							
F:	2-610/Product:	glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted							
F:	2/Active site:	Cys #status predicted							

Query Match 72.9%; Score 2250.5; DB 2; Length 610;
Best Local Similarity 71.3%; Pred. No. 5.4e-138; Mismatches 102; Indels 1; Gaps 1;
Matches 435; Conservative 77;

QY 1 MCGIVGAIAQDVAEILLEGRLREYRGYSAGLAVVDTTEGHMTRRLRLKGVQMLAQAAE 60
DB 1 MCGIVGAVAGDVAEILVQGLRLREYRGYSAGVAVVDSKQLTRRLKGVQELADAVE 60

QY 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHV-SHIVVHNGIIEHHEPLEELKARGYTFV 120
DB 61 AAQVAGCTGTAHTRWATHGEPSEVNAHPHISGDIIVVHNGIIEHHEMLRMLDQGVFT 120

QY 121 SETDTEVIAHLVNMELKOGGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
DB 121 SQTDETVIAHLVNMELKOGGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

QY 181 IGLGMENFIASDQALLPVTFRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
DB 181 IGFIGENFLASDQALLPVTFRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYCHYVWKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTOILA 300
DB 241 HDAADKGYRHYWKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTOILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLSQSGETADTLA 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLSQSGETADTLA 360

QY 361 RLKSELGYLGSIAICNVPGSSLVRESLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
DB 361 RLAKENGYMAAMTICNVAGSSLVRESLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420

QY 421 SR-LKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEFSDKHALFLSRGQOYPI 479
DB 421 GKQOQIRIGRELAIEIHALQPKQIBETALSFEKQIETLAEDFADKHHTFLGRGEYPI 480

QY 480 ALLEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEBVR 539
DB 481 AVASLKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEBVR 540

QY 540 RGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQ 599
DB 541 RGGLYVFADEVAGFEADTKMIAMPHVSEIAPYITPTMQLLSYHVALIKGTVDQ 600

QY 600 RNLAKEVTVVE 609
DB 601 RNLAKEVTVVE 610

RESULT 8
F82951
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - F
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F82951
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F82951
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-611 <STO>
A;Cross-references: UNIPROT:Q9HT25; UNIPARC:UPI0000165E0; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: glms; PA5549
C;Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-611/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
F;2/Active site: Cys #status predicted

Query Match 64.2%; Score 1983; DB 2; Length 611;
Best Local Similarity 62.4%; Pred. No. 1.2e-120; Mismatches 127; Indels 2; Gaps 2;
Matches 381; Conservative 101;

QY 1 MCGIVGAIAQDVAEILLEGRLREYRGYSAGLAVVDTTEGHMTRRLRLKGVQMLAQAAE 60
DB 1 MCGIVGAIAERNITPILIEGLKLEVRGYSAGVAVFDEGRQRCRVKVASLEBGLA 60

QY 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHV-SHIVVHNGIIEHHEPLEELKARGYTF 119
DB 61 GTPLGLGLGTAHTRWATHGEPSEVNAHPHSSDEVAHVHNGIIEHHEPLEERLKLGLYVF 120

QY 120 VSETDTEVIAHLVNMELKOGGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPL 179
DB 121 TSQDTEVIAHLVNMELKOGGTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPL 180

QY 180 VIGLGMENFIASDQALLPVTFRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNL 239
DB 181 VIGLGMENFLASDQALLPVTFRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNL 240

QY 240 QYDAGDKGIYCHYVWKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTOILA 299
DB 241 GAEADKGYRHYWKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTOILA 300

QY 300 ACCTSNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLSQSGETADTLA 359
DB 301 ACCTSNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLSQSGETADTLA 360

QY 360 LRSLKSELGYLGSIAICNVPGSSLVRESLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAK 419
DB 361 LRNAKSELGFLSSVAICNVATSSLVRESLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAK 420

QY 420 LSRL-KGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEFSDKHALFLSRGQOY 478
DB 421 IGQVQRLGADGVBAELVDELRLPTLGEALAMNRTVEKVSSELFAEKHHTFLGRGAQFP 480

QY 479 IALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEBVR 538
DB 481 VALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEBVR 540

QY 539 ARGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQ 598
DB 541 ARGGLYVFADEVAGFEADTKMIAMPHVSEIAPYITPTMQLLSYHVALIKGTVDQ 600

QY 599 PRNLAKEVTVVE 609
DB 601 PRNLAKEVTVVE 611

RESULT 9
A84933
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Buch
N;Alternate names: D-fructose-6-phosphate amidotransferase
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 28-Jul-2003
C;Accession: A84933
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: A84933
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-621 <STO>
A;Cross-references: UNIPARC:UPI000005E424; GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: glms; BU026
C;Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
Query Match 60.1%; Score 1856; DB 2; Length 621;

		Best Local Similarity 58.8%; Pred. No. 28-112; Matches 359; Conservative 116; Mismatches 132; Indels 4; Gaps 3;			
Qy	1	MCGIAGTAQRDVAEILLEGRLRYGYDSAGLAVVDTTEGHMTRLRLRLGKVMQLAAQAE	60		
Db	13	MCGIAGTAQRDVAEILLEGRLRYGYDSAGLAVVDTTEGHMTRLRLRLGKVMQLAAQAE	60		
Qy	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV	120		
Db	73	KKKILGSGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV	120		
Qy	121	SETDEVIAHLVNWEL-KGGTTLRAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPL	179		
Db	133	SDTDEVIAHLVNWEL-KGGTTLRAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPL	179		
Qy	180	VIGLNGENFIASDQALLPVTRRFIFLEGDIAETTRSVNIFDKTGAEVKRODIESNL	239		
Db	193	IIGLNGENFIASDQALLPVTRRFIFLEGDIAETTRSVNIFDKTGAEVKRODIESNL	239		
Qy	240	QYDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISH-GQVDLSLSELGPNADLLSKVEHIQI	298		
Db	253	EYSAKKGKRYRMEKEIHEQPKSRINTLKNRLTNSKNVHFSELGSKENNIFYNTEHIQI	312		
Qy	299	LACGTSYNSGMSYRWYFESLAGIPCDVEIASFRYKSAVRNSLMITLSQSGETADTLA	358		
Db	313	VACGTSYNSGMSYRWYFESLAGIPCDVEIASFRYKSAVRNSLMITLSQSGETADTLA	358		
Qy	359	GURLSKELGVLGLSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPTQTTLVLLMLVA	418		
Db	373	ALRYSKGLGVLGLSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPTQTTLVLLMLVA	418		
Qy	419	KLSRLKGLDASTEHIDVHGLQALPSRIEQMLSDQKRIEALAEFDSKHALPLSRGQYVP	478		
Db	433	KIINGSKENNTSKRIVQTLSPVRIEILKQKLIQDMANTLANKNMFLGRGNQYP	492		
Qy	479	ILEGALKLKETSITHABAYAGELKHGPLALIDADMPVIVVAPNNLEKLKSNIEEVR	538		
Db	493	IAMEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLEKLKSNIEEVR	538		
Qy	539	ARGGOLYFADQDQGVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDQ	598		
Db	553	SRGGIVYVPSNQE--FDYENINTIKLPYVEELIAPFYTVPLQLLAYHVALIKGTDVDQ	610		
Qy	599	PRNLAKSVTVE 609			
Db	611	PHRLAKSVTVE 621			

RESULT 10
E1272
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms - syphilis
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: E71272
R:Frazer, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: E71272
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-635 <COL>
A:Cross-references: UNIPROT:O83833; UNIPARC:UPI0000164AEC; GB:AE001256; GB:AE000520; NID
A:Experimental source: strain Nichols
A:Genetics:
C:Gene: TP0861
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-635/Product: Glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
F:2/Active site: Cys #status predicted

		Query Match 54.6%; Score 1688; DB 2; Length 635; Best Local Similarity 53.7%; Pred. No. 1.6e-101; Matches 341; Conservative 98; Mismatches 170; Indels 26; Gaps 5;			
Qy	1	MCGIAGTAQRDVAEILLEGRLRYGYDSAGLAVVDTTEGHMTRLRLRLGKVMQLAAQAE	60		
Db	1	MCGIAGTAQRDVAEILLEGRLRYGYDSAGLAVVDTTEGHMTRLRLRLGKVMQLAAQAE	60		
Qy	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV	120		
Db	61	QSPCLGCTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV	120		
Qy	121	SETDEVIAHLVNWELKGGTTLRAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV	180		
Db	121	SQTDEVIAHLVNWELKGGTTLRAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV	180		
Qy	181	IGLNGENFIASDQALLPVTRRFIFLEGDIAETTRSVNIFDKTGAEVKRODIESNL	240		
Db	181	VGLNGENFIASDQALLPVTRRFIFLEGDIAETTRSVNIFDKTGAEVKRODIESNL	240		
Qy	241	YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGQVDLSL-----GPN-----D	287		
Db	241	LCTQDKGTHRHMHQEIWQPHAIRHTLNAVYFSSSSRAQVTRFGEDRVLDGTSCKTPE	300		
Qy	288	ELLSKVEHIQILACCTSYNSGMSYRWYFESLAGIPCDVEIASFRYKSAVRNSLMITL	347		
Db	301	RLFRITRVRITACCTSYHAGLVARYWFEAFAGVCQVEIASFRYKSAVRNSLMITL	360		
Qy	348	SQSGTADTLAGRLSKELGVLGLSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPT	407		
Db	361	SQSGTADTLAGRLSKELGVLGLSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPT	407		
Qy	408	TQTLVLLMLVAKLSRLKGLDASTEHIDVHGLQALPSRIEQMLSDQKRIEALAEFDSKH	466		
Db	421	TQTLVLLMLVAKLSRLKGLDASTEHIDVHGLQALPSRIEQMLSDQKRIEALAEFDSKH	466		
Qy	467	HALFLSRGDOVPALLEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNIE	525		
Db	481	HALFLSRGDOVPALLEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNIE	525		
Qy	526	-LLEKLKSNIEEVRARGGQLYFAD-----QDAGFVSSDNNHIIEMPHVEVIAPI	574		
Db	541	VLFEKASNIEEVRARGGQLYFAD-----QDAGFVSSDNNHIIEMPHVEVIAPI	600		
Qy	575	IFYTVPLQLLAYHVALIKGTDVDQPRNLAKSVTVE 609			
Db	601	IFYAVPLQLLAYHVALIKGTDVDQPRNLAKSVTVE 635			

RESULT 11
A82844
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - X
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82844
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
ce 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82844
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-609 <SIM>
A:Cross-references: UNIPROT:Q9PH05; UNIPARC:UPI0000165A4A; GB:AE003867; GB:AE003849; NID:
A:Experimental source: strain 95C
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.

A;Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.P.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, F.A.; da Silva, W.A.; da Silveira M.; Tsuchako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0141
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
P;2-609/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F;2/Active site: Cys #status predicted

```
Query Match 54.4%; Score 1681; DB 2; Length 609;
Best Local Similarity 54.6%; Pred. No. 4.2e-101;
Matches 333; Conservative 103; Mismatches 172; Indels 2; Gaps 2;

Qy 1 MCGIVGAIAQDVAEIILEGRLRLEYRGYSAGLAVVDTEGHMTPLRLRLGKVMQLAQAAE 60
Db 1 MCGIVGAIAQDVAEIILEGRLRLEYRGYSAGLAVVDTEGHMTPLRLRLGKVMQLAQAAE 60

Qy 61 EPHLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTFV 120

Qy 60 QEGFTASLGIGHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTFV 119
Db 60 QEGFTASLGIGHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTFV 119

Qy 121 SETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

Qy 120 SQDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 179
Db 120 SQDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 179

Qy 181 IGLMGSENFASDQALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNLQ 240
Db 181 IGLMGSENFASDQALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNLQ 240

Qy 180 IGIDGHELVASDISAVIQATQVIFLEDGDTAIBRDGSIFFNAEQCPVERPLHLSNV 239
Db 180 IGIDGHELVASDISAVIQATQVIFLEDGDTAIBRDGSIFFNAEQCPVERPLHLSNV 239

Qy 241 YDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSELGNPADELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSELGNPADELLSKVEHIQILA 300

Qy 240 LSSLELGEFRHFQKEIHEQPRVLADTWEAIDAAGPPMLFGAQAESVFRGITGIQILA 299
Db 240 LSSLELGEFRHFQKEIHEQPRVLADTWEAIDAAGPPMLFGAQAESVFRGITGIQILA 299

Qy 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASEFPRKSAVRNLSMTITLSOSGETADTLA 360
Db 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASEFPRKSAVRNLSMTITLSOSGETADTLA 360

Qy 300 CGTSYAGLARYWIEAIGLPCHEIASEYRYKAYVNPQHVLVTTISOSGETLDTLEAL 359
Db 300 CGTSYAGLARYWIEAIGLPCHEIASEYRYKAYVNPQHVLVTTISOSGETLDTLEAL 359

Qy 361 RLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420
Db 361 RLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420

Qy 360 KYAKALGHRHTLSICNAPDSAIPRISLICYTRAGPEIGVASTKAFPTTQTLVLLMLVAKL 419
Db 360 KYAKALGHRHTLSICNAPDSAIPRISLICYTRAGPEIGVASTKAFPTTQTLVLLMLVAKL 419

Qy 421 SRLKG-LDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRGDQYPI 479
Db 421 SRLKG-LDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRGDQYPI 479

Qy 420 GVLRGAVDAEHAAYLEQLRQLPCGVQOALNLEPQIAAWAECFASRHHALFLGRGLHVP 479
Db 420 GVLRGAVDAEHAAYLEQLRQLPCGVQOALNLEPQIAAWAECFASRHHALFLGRGLHVP 479

Qy 480 ALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSNIEVRA 539
Db 480 ALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSNIEVRA 539

Qy 480 ALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSNIEVRA 539
Db 480 ALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSNIEVRA 539

Qy 540 RGGOLYVFADQDAGFVSDNNHIEEMPHVEVIAPIETVPLQLLAYHVALIKCTDQVDP 599
Db 540 RGGOLYVFADQDAGFVSDNNHIEEMPHVEVIAPIETVPLQLLAYHVALIKCTDQVDP 599

Qy 540 RGGELFVFADQDHFSESEGLHVIRTLRHGTGLVSLVHTTIPVQLLAYHVALIKCTDQVDP 599
Db 540 RGGELFVFADQDHFSESEGLHVIRTLRHGTGLVSLVHTTIPVQLLAYHVALIKCTDQVDP 599

Qy 600 RNLAKSVTVE 609
Db 600 RNLAKSVTVE 609
```

RESULT 12
T45493
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms [similarity
N;Alternate names: Glucosamine-6-phosphate synthase
C;Species: Thibacillus ferrooxidans
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45493
R;Oppon, J.C.; Sarnovsky, R.J.; Craig, N.L.; Rawlings, D.E.
J. Bacteriol. 180, 3007-3012, 1998
A;Title: A Tn7-like transposon is present in the gimUS region of the obligately chemoaut

A;Reference number: Z22992; MUID:98269023; PMID:9603897
A;Accession: T45493
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-611 <OPP>
A;Cross-references: UNIPROT:Q56275; UNIPARC:UPI000016E22C; EMBL:AF032884; NID:g2653994;
A;Experimental source: ATCC 33020
C;Genetics:
A;Gene: glms
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase

```
Query Match 51.1%; Score 1579; DB 2; Length 611;
Best Local Similarity 52.4%; Pred. No. 1.7e-94;
Matches 322; Conservative 101; Mismatches 182; Indels 10; Gaps 5;

Qy 1 MCGIVGAIAQDVAEIILEGRLRLEYRGYSAGLAVVDTEGHMTPLRLRLGKVMQLAQAAE 60
Db 1 MCGIVGVSKTDLVPMILEGLQRLEYRGYSAGLAILGADADALLRVRSGRVAELTAADV 60

Qy 61 EPHLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTF 119
Db 61 EPHLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTF 119

Qy 61 ERGLQGVGIGHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTF 120
Db 61 ERGLQGVGIGHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREELKARGYTF 120

Qy 120 VSETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPL 179
Db 120 VSETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPL 179

Qy 121 TSETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPL 180
Db 121 TSETDTEVIAHLVNMWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPL 180

Qy 180 VIGLMGSENFASDQALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNL 239
Db 180 VIGLMGSENFASDQALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNL 239

Qy 181 LLGVADDDGHYFASDVAALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNL 240
Db 181 LLGVADDDGHYFASDVAALLPVTRRFLFLEGDIATERRSVNIPDKTGAEVKRODIESNL 240

Qy 240 QYDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSELGNPADELLSKVEHIQI 298
Db 240 QYDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSELGNPADELLSKVEHIQI 298

Qy 241 SAAAVDLGPRYHFMQKEIHEQPRAVADTLEGAL-NSQLDLDLMDGGAAMFRDVRVLF 299
Db 241 SAAAVDLGPRYHFMQKEIHEQPRAVADTLEGAL-NSQLDLDLMDGGAAMFRDVRVLF 299

Qy 299 LACGTSYNSGMVSRWYFESLAGIPCDVEIASEFPRKSAVRNLSMTITLSOSGETADTLA 358
Db 299 LACGTSYNSGMVSRWYFESLAGIPCDVEIASEFPRKSAVRNLSMTITLSOSGETADTLA 358

Qy 300 LAGTSHYATLVGQWVESIVGIPQAQELGHEYRYRDSIPDPQOLVLTLSOSGETLDTFE 359
Db 300 LAGTSHYATLVGQWVESIVGIPQAQELGHEYRYRDSIPDPQOLVLTLSOSGETLDTFE 359

Qy 359 GLRLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVA 418
Db 359 GLRLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVA 418

Qy 360 ALRRKDLGHTRTLAI CNVAESAIPRASALRFLFRAGPEIGVASTKAFPTTQTLVLLMLVA 419
Db 360 ALRRKDLGHTRTLAI CNVAESAIPRASALRFLFRAGPEIGVASTKAFPTTQTLVLLMLVA 419

Qy 419 KLSRLKGLDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRG 474
Db 419 KLSRLKGLDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRG 474

Qy 420 SLAKAPG---ASERCAAGSGPGRQLRQLPGSVQOALNLEPQIAAWAECFASRHHALFLGRG 476
Db 420 SLAKAPG---ASERCAAGSGPGRQLRQLPGSVQOALNLEPQIAAWAECFASRHHALFLGRG 476

Qy 475 DOYPIALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSN 534
Db 475 DOYPIALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSN 534

Qy 477 LHYPIALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSN 536
Db 477 LHYPIALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELEKLSN 536

Qy 535 EYVRARGOLYVFADQDAGFVSDNNHIEEMPHVEVIAPIETVPLQLLAYHVALIKCTDQV 594
Db 535 EYVRARGOLYVFADQDAGFVSDNNHIEEMPHVEVIAPIETVPLQLLAYHVALIKCTDQV 594

Qy 537 QEVHARGELYVFADSDSHFNAGSVHVMRLPRHAGLLSPVHAIPVQLLAYHVALIKCTDQV 596
Db 537 QEVHARGELYVFADSDSHFNAGSVHVMRLPRHAGLLSPVHAIPVQLLAYHVALIKCTDQV 596

Qy 595 DVDPQPNLAKSVTVE 609
Db 595 DVDPQPNLAKSVTVE 609
```

RESULT 13
H82022
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) NMA0276 [similar
N;Alternate names: Glucosamine fructose-6-phosphate aminotransferase [mnsomer]
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: H82022
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitens*
A:Reference number: AD3252; PMID:1175668
A:Accession: AD3595
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <KUR>
A:Cross-references: UNIPROT:Q8YC47; UNIPARC:UPI000005859C; GB:AE008918; PIDN:AAL53927.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10685
A:Map position: II
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase

Query Match		46.5%;	Score 1435;	DB 2;	Length 612;
Best Local Similarity		48.9%;	Pred. No. 3.7e-85;		
Matches 300;		Conservative 99;	Mismatches 204;	Indels 10;	Gaps 6;
Qy	1	MCGIVGAIAQORDVAEII	LEGLRLLEYRGYSAGLAVVDTEGHMTLRRLGKVMQAQAE	60	
Db	6	MCGIIGIIGNDEVAPRLV	DALKRLLEYRGYSAGIATIQ-NGRLDRRAEGKLVNLEKRLA	64	
Qy	61	EHLHGCTGIAHTWATHGP	SEVNAPHVSEHIVVHNGIIENHEPLREBELKARGYTFV	120	
Db	65	GEPLPGVIGICHTWATHG	RPEVHNAPHITTRLAVVHNGIIENFAELRAMLEAEGKFE	124	
Qy	121	SETDTEVIAHLVNMELKQ	GTLEAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV	180	
Db	125	TETDTEAVHLVTRLEKSK	SPVEAVRDCLPHLKGAFAFALFPEGDELLIGARQGPPLA	184	
Qy	181	IGLGMENFTASDQALLP	PTRRFIFLEEGDIAEITRRSVNIPDKTAEVKRQDIESNLQ	240	
Db	185	VGYGEGEMFLGSDAIAL	APFTDTISYLEDGDWAVLTRNGVSIYDENNKPVERPQKSQNT	244	
Qy	241	YDAGDKGIYCHYMOKEI	YEOFPNAIKNTLTGRI--SHGOVDLSLGPNADELLSKVEHIQI	298	
Db	245	NMLVSKGNHRHFMQEN	PEQEVISHLTANYLDFTTKVRKEAIGID----FSKVDRLTI	300	
Qy	299	LACGTSYNSGMVSRWFES	LAGIPCDVEIASSEFYRKSAVRRNSLMITLSQSGETADTLA	358	
Db	301	TACCTAYYAATVAKYWF	EQIARLPVDSDIASEFYREMPKSLKSLAMFVSQSGETADTLA	360	
Qy	359	GLRLSKELGYLGLSLAI	CNVPSSLVRESDLALMTNAGTEIGVASTKAFDTQTLVLLMLVA	418	
Db	361	SLRYCKAQG-LKTASVL	NVTGSTIARESDAVFFLAGFEIGVASTKAFDTQTLVLLMLVA	419	
Qy	419	KLRLKG-LDASIEHDI	VHGLQALPSRIEOMLSODKRIEALAEDEFSDKHAFILSRGDQY	477	
Db	420	AAARAGAIDEVREQEL	VHQLSEAPRFINQVLEQDIADVCHDLKVNHVLYLGRGTSF	479	
Qy	478	PIALEGALKKEISYTHA	EYAGELKHGPIALIDAMPVIVVAPNNLEKLSNIEEV	537	
Db	480	PLAMEGALKKEISYTHA	EYAGELKHGPIALIDETMPVIVVAPSRLYEKTVSNMQEV	539	
Qy	538	RARGQLYVPAD-QDAGF	VSDNMHIIEMPHVEVIAPIFYTPVQLLAYHVALIKGTDV	596	
Db	540	AARGGRIILITDRKGA	BSASIDTWTATVLPVPEPFISPLVYALPIQMLAYHTAVLMGTDV	599	
Qy	597	DQPNLAKSVTVE	609		
Db	600	DQPNLAKSVTVE	612		

Search completed: June 14, 2006, 15:28:37
Job time : 43.2344 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:14:44 ; Search time 74.7383 Seconds
(without alignments)
972.693 Million cell updates/sec

Title: US-10-612-779-30
Perfect score: 835
Sequence: 1 MSLLPQGFYIRMEEGDLEQV.....VKFYKCGFSNAGVEMQIRK 159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*

- 1: geneseqp1980s:*
- 2: geneseqp1390s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*
- 10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	100.0	159	ADI38885	Adi38885 Glucosami
2	835	100.0	159	ADJ64176	Adj64176 S. cerevi
3	835	100.0	159	ADU00600	Adu00600 Amino aci
4	333	39.9	149	ABP73464	Abp73464 Candida a
5	328	39.3	149	ADI38887	Adi38887 Glucosami
6	328	39.3	149	ADU00602	Adu00602 Amino aci
7	275	32.9	190	ABJ26242	Abj26242 Aspergill
8	261.5	31.3	149	ADI38889	Adi38889 Glucosami
9	216.5	25.9	174	ADY05609	Ady05609 Plant ful
10	209	25.0	184	AAU07779	Aau07779 Human nov
11	208	24.9	184	AAU73505	Aau73505 Human tra
12	208	24.9	184	AAU67123	Aau67123 Amino aci
13	208	24.9	184	ADRI0847	Adri0847 Human FLJ
14	208	24.9	184	ADRI0848	Adri0848 Human FLJ
15	208	24.9	200	AAU63700	Aau63700 Human gas
16	207	24.8	219	ABBS8239	Abbs8239 Drosophil
17	204	24.4	184	ABU65053	Abu65053 Human NOV
18	203	24.3	206	AAU79992	Aau79992 Human pro
19	202.5	24.3	180	ABJ25642	Abj25642 Aspergill
20	202	24.2	184	ABU65054	Abu65054 Human NOV
21	202	24.2	184	ADN61759	Adn61759 Human nov
22	199	23.8	184	AAU79008	Aau79008 Human pro
23	189.5	22.7	183	ADN61757	Adn61757 Human nov

24	158	18.9	107	4	AAU23427	Novel hum
25	126	15.1	154	5	ABP27501	Streptoco
26	126	15.1	154	8	ADV81684	Streptoco
27	126	15.1	156	8	ADV88261	Streptoco
28	126	15.1	156	8	ADV79514	Streptoco
29	117.5	14.1	242	7	ABO79472	Pseudomon
30	116.5	14.0	157	5	ABB50043	Listeria
31	113	13.5	154	5	ABB53909	Lactococc
32	112	13.4	140	6	ABU23792	Protein e
33	111	13.3	166	6	ABM72742	Staphyloc
34	110	13.2	179	4	AAU78880	C. glutam
35	110	13.2	179	4	AAU90861	C. glutam
36	106	12.7	149	7	ADC95669	E. faeciu
37	106	12.7	185	5	ABB47559	Abb47559 Listeria
38	106	12.7	306	6	ABM70329	Abm70329 Photorhab
39	105.5	12.6	159	5	AAE25024	AAE25024 Human dru
40	104	12.5	307	7	ADF06564	ADF06564 Bacterial
41	102	12.2	154	9	ABE42282	L. pneumo
42	102	12.2	184	9	ABE36279	ABE36279 L. pneumo
43	101	12.1	134	8	ADP30044	ADP30044 Human sec
44	101	12.1	168	5	ABB49030	ABB49030 Listeria
45	101	12.1	168	6	ABU32907	Abu32907 Protein e

ALIGNMENTS

RESULT 1
ADI38885
ID ADI38885 standard; protein; 159 AA.
XX
AC ADI38885;
XX
DT 15-APR-2004 (first entry)
XX
DE Glucosamine-6-phosphate acetyltransferase, GNA1, SEQ ID 30.
XX
KW Glucosamine; N-acetylglucosamine; fermentation;
KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.
XX
OS Saccharomyces cerevisiae.
XX
PN WO2004003175-A2.
XX
PD 08-JAN-2004.
XX
PF 01-JUL-2003; 2003WO-US020925.
XX
PR 01-JUL-2002; 2002US-0393348P.
XX
PA (ARKI-) ARKION LIFE SCI LLC.
XX
PI Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
DR WPI; 2004-203380/19.
XX
N-PSDB; ADI38884.
XX
PT Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
PS Claim 5; SEQ ID NO 30; 327pp; English.
XX
CC The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate

CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
 CC the activity of (IV) and increases the activity of glucosamine-1
 CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
 CC which is chosen from the group consisting of glucosamine-6-phosphate, N-
 CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
 CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
 CC sequence was used to illustrate the method of the invention.
 XX
 SQ Sequence 159 AA;
 Query Match 100.0%; Score 835; DB 8; Length 159;
 Best Local Similarity 100.0%; Pred. No. 3e-83;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLPDGFYIRMEEDLEQVETTLKVLTVGTITPESFSKLIKYWNEATVWVNDKIM 60
 DB 1 MSLPDGFYIRMEEDLEQVETTLKVLTVGTITPESFSKLIKYWNEATVWVNDKIM 60
 QY 61 QYNPMVIVDKRTETVAATGNIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDOLV 120
 DB 61 QYNPMVIVDKRTETVAATGNIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDOLV 120
 QY 121 TIGFDYGCYKIILDCDEKNVFKYECGFSNAGVEMQIRK 159
 DB 121 TIGFDYGCYKIILDCDEKNVFKYECGFSNAGVEMQIRK 159
 RESULT 2
 ADJ64176
 ID ADJ64176 standard; protein; 159 AA.
 XX
 AC ADJ64176;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE S. cerevisiae glucosamine phosphate N-acetyltransferase.
 XX
 KW Hexosamine; arthritis; osteoarthritis; rheumatoid arthritis; pain;
 KW inflammation; articular joint disease; spondyloarthritis;
 KW gouty arthritis; systemic lupus erythematosus; juvenile arthritis;
 KW tendinitis; bursitis; connective tissue injury; psoriasis; eczema; burn;
 KW dermatitis; transgenic plant; glucosamine phosphate N-acetyltransferase;
 KW yeast; enzyme; EC 2.3.1.4.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US2004003432-A1.
 XX
 PD 01-JAN-2004.
 XX
 PF 05-MAY-2003; 2003US-00429812.
 XX
 PR 06-MAY-2002; 2002US-0378297P.
 XX
 PA (PHAA) PHARMACIA CORP.
 XX
 PI Obukowicz MG;
 XX
 DR WPI; 2004-061704/06.
 DR N-PSDB; ADJ64175.
 XX
 PT Producing hexosamine useful for treating arthritis by culturing cell
 PT comprising heterologous polynucleotide sequences which encode for enzymes
 PT required for biosynthetic pathway capable of synthesizing hexosamine.
 XX
 PS Disclosure; SEQ ID NO 4; 35pp; English.
 XX
 CC The invention relates to a method of producing a hexosamine that involves
 CC providing a cell comprising polynucleotide sequences which code for
 CC enzymes required for a biosynthetic pathway capable of synthesizing the
 CC hexosamine, where at least one of the polynucleotide sequences comprises
 CC a recombinant polynucleotide. The method involves transforming a cell
 CC with at least one heterologous polynucleotide coding for a polypeptide in

CC a biosynthetic pathway that is capable of producing hexosamine, culturing
 CC the transformed cell under conditions that permit the cell to translate
 CC the polynucleotide into a polypeptide comprising an enzyme which is part
 CC of the biosynthetic pathway. The hexosamine produced by the method of
 CC invention is useful for treating arthritis, preferably osteoarthritis and
 CC rheumatoid arthritis in humans and pet animals. The hexosamines are also
 CC useful for treating pain, inflammation or inflammation-associated
 CC disorder. The hexosamines are also useful for treating other articular
 CC joint damage or diseases such as spondyloarthritis, gouty arthritis,
 CC systemic lupus erythematosus and juvenile arthritis and other joint
 CC damage disease conditions such as tendinitis, bursitis, connective tissue
 CC injuries or disorders, and skin related conditions such as psoriasis,
 CC eczema, burns and dermatitis. The method provides improved production of
 CC hexosamine without compromising its clinical effectiveness and in a form
 CC that is acceptable for delivery to a broad class of patients. Consumption
 CC of edible portions of transgenic plants containing hexosamines provide a
 CC supply of hexosamine to the subject in a conventional and easily
 CC consumable form. The present sequence is Saccharomyces cerevisiae
 CC glucosamine phosphate N-acetyltransferase (EC 2.3.1.4).
 XX
 SQ Sequence 159 AA;
 Query Match 100.0%; Score 835; DB 8; Length 159;
 Best Local Similarity 100.0%; Pred. No. 3e-83;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLPDGFYIRMEEDLEQVETTLKVLTVGTITPESFSKLIKYWNEATVWVNDKIM 60
 DB 1 MSLPDGFYIRMEEDLEQVETTLKVLTVGTITPESFSKLIKYWNEATVWVNDKIM 60
 QY 61 QYNPMVIVDKRTETVAATGNIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDOLV 120
 DB 61 QYNPMVIVDKRTETVAATGNIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDOLV 120
 QY 121 TIGFDYGCYKIILDCDEKNVFKYECGFSNAGVEMQIRK 159
 DB 121 TIGFDYGCYKIILDCDEKNVFKYECGFSNAGVEMQIRK 159
 RESULT 3
 ADU00600
 ID ADU00600 standard; protein; 159 AA.
 XX
 AC ADU00600;
 XX
 DT 13-JAN-2005 (first entry)
 XX
 DE Amino acid sequence of glucosamine-6-phosphate acetyltransferase.
 XX
 KW chitin; chitosan; fermentation;
 KW glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;
 KW glucosamine-6-phosphate synthase; GFA1 gene; glms gene;
 KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;
 KW chitin deacetylase; CDA1 gene; CDA2 gene;
 KW N-acetylglucosamine-6-phosphate deacetylase;
 KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;
 KW enzyme.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO2004092391-A2.
 XX
 PD 28-OCT-2004.
 XX
 PF 12-APR-2004; 2004WO-US011286.
 XX
 PR 11-APR-2003; 2003US-0462087P.
 XX
 PA (ARKI-) ARKION LIFE SCI LLC.
 XX
 PI Deng M, McMullin TW, Grund AD;
 XX WPI; 2004-766880/75.
 DR

DR N-PSDB; ADU00599.

XX Producing chitin or chitosan, involves culturing microorganism with

PT genetic modification that results in increase in activity of glutamine-

PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce

PT chitin or chitosan.

XX Example 1; SEQ ID NO 33; 161pp; English.

XX The specification describes a method for producing chitin or chitosan by

CC a fermentation process. The method involves culturing in a fermentation

CC medium a microorganism which comprises one or more genetic modifications

CC that result in an increase in the activity of glutamine-fructose-6-

CC phosphate amidotransferase (also known as glucosamine synthetase and

CC glucosamine-6-phosphate synthase), and encoded by the GNA1 eukaryotic gene

CC and the gms bacterial gene), glucosamine-6-phosphate acetyltransferase

CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1

CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-

CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and

CC chitosanase, and collecting the chitin or chitosan. The method is useful

CC for producing chitin or chitosan by utilizing microorganisms such as

CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and

CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.

CC cerevisiae, A. niger or A. nidulans. The method enables high quantities

CC of chitin and chitosan to be produced cost effectively. The present

CC sequence is encoded by a GNA1 gene, and is a glucosamine-6-phosphate

CC acetyltransferase. The GNA1 gene was used to transform yeast for use in

CC the method of the invention.

XX Sequence 159 AA;

XX Query Match 100.0%; Score 835; DB 8; Length 159;

XX Best Local Similarity 100.0%; Pred. No. 3e-83;

XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLPDGYFIRRMEEGDLEQVTTTLKVLTVGTITPESFSLIKYWNATVWMDNEDKKIM 60

DB 1 MSLPDGYFIRRMEEGDLEQVTTTLKVLTVGTITPESFSLIKYWNATVWMDNEDKKIM 60

QY 61 QYNPMVIVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLIDQLV 120

DB 61 QYNPMVIVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLIDQLV 120

QY 121 TIGFDYGCYKIIILDCDEKNVKEKCFGSNAGVEMQIRK 159

DB 121 TIGFDYGCYKIIILDCDEKNVKEKCFGSNAGVEMQIRK 159

RESULT 4

ABP73464

ID ABP73464 standard; protein; 149 AA.

AC ABP73464;

XX 30-JAN-2003 (first entry)

DE Candida albicans essential protein SEQ ID NO 7301.

XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;

KW signal transduction; DNA replication; cell division; growth;

KW proliferation; Candida albicans; fungicide; antifungal.

XX Candida albicans.

OS WO200253728-A2.

XX 11-JUL-2002.

XX 26-DEC-2001; 2001WO-0549486.

XX 29-DEC-2000; 2000US-0259128P.

PR 20-FEB-2001; 2001US-00792024.

PR 22-AUG-2001; 2001US-0314050P.

XX (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

PI WPI; 2002-566694/60.

XX N-PSDB; ABZ32014.

XX Constructing strains for identifying gene products as effective targets

PT for therapeutic intervention, by inactivating in the strain one allele of

PT a gene and placing other allele of the gene under conditional expression.

XX Claim 44; SEQ ID NO 7301; 167pp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal

CC cells in which both alleles of a gene are modified, comprising modifying

CC one allele by insertion or replacement by a cassette having an

CC expressible selectable marker and modifying other allele by

CC recombination, of a promoter replacement fragment with a heterologous

CC promoter, so that expression of the second allele is regulated by the

CC promoter. (M1) is useful for constructing a strain of diploid fungal

CC cells in which both alleles of a gene are modified. The diploid fungal

CC cells having both alleles modified are useful for identifying a gene that

CC is essential to the survival or growth of a fungus, a gene that

CC contributes to the virulence and/or pathogenicity of a fungus, a gene

CC that contributes to the resistance of a diploid fungus to an antifungal

CC agent, an antifungal agent that inhibits the growth of a diploid fungus

CC and for identifying a therapeutic agent for treatment of a mammalian

CC disease. (M1) is useful for identifying a compound which modulates the

CC activity of a gene product, preferably enzymatic activity, carbon

CC compound catabolism, biosynthetic, transporter, transcriptional,

CC translational, signal transduction, DNA replication and cell division

CC activity. The method is useful for identifying a compound having the

CC ability to inhibit growth or proliferation of C. albicans cells and for

CC treating infection by C. albicans. The present sequence is that of an

CC essential Candida albicans protein used in the method of the invention.

CC Note: The sequence data for this patent is not represented in the printed

CC specification but is based on sequence information supplied to Derwent by

CC the European Patent Office

XX Sequence 149 AA;

XX Query Match 39.9%; Score 333; DB 5; Length 149;

XX Best Local Similarity 44.0%; Pred. No. 4.2e-28;

XX Matches 70; Conservative 26; Mismatches 51; Indels 12; Gaps 3;

QY 1 MSLPDGYFIRRMEEGDLE-QVTETLKVLTVTGTITPESFSLIKYWNATVWMDNEDKKI 59

DB 1 MSLPDGYFIRRMEEGDLE-QVTETLKVLTVTGTITPESFSLIKYWNATVWMDNEDKKI 59

QY 60 MOYNPMVIVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLIDQLV 119

DB 51 SIYHPYVITN-ASGIWVATGMLFVEKKLIHECGKVGHTIEDISVAKSEQGKGLGYLVTSL 109

QY 120 VTIGFDYGCYKIIILDCDEKNVKEKCFGSNAGVEMQIR 158

DB 110 TKVAQENDCYKILDCSPENGVGYEKGKGYKDGVEVNCR 148

RESULT 5

AD138887

ID AD138887 standard; protein; 149 AA.

XX AD138887;

XX 15-APR-2004 (first entry)

XX Glucosamine-6-phosphate acetyltransferase, GNA1, SEQ ID 32.

XX Glucosamine; N-acetylglucosamine; fermentation;

KW glucosamine-6-phosphate acetyltransferase;

KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;

KW glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;

KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.
XX Candida albicans.
OS WO2004003175-A2.
PN 08-JAN-2004.
PD 01-JUL-2003; 2003WO-US020925.
PF 01-JUL-2002; 2002US-0393348P.
PR (ARKI-) ARKION LIFE SCI LLC.
PA Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX WPI; 2004-203380/19.
DR N-PSDB; ADI38886.
XX Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
XX Claim 5; SEQ ID NO 32; 327pp; English.
XX The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate
CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
CC the activity of (IV) and increases the activity of glucosamine-1
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
CC which is chosen from the group consisting of glucosamine-6-phosphate,
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
CC sequence was used to illustrate the method of the invention.
XX Sequence 149 AA;
Query Match 39.3%; Score 328; DB 8; Length 149;
Best Local Similarity 43.4%; Pred. No. 1.5e-27;
Matches 69; Conservative 26; Mismatches 52; Indels 12; Gaps 3;
QY 1 MSLPDGFYIRMEEGDLE-QVTEILKVLTTVTGTTTSPESFKLIKYWNEATVWNNEDKKI 59
DB 1 MMLPGQYTFRKLKLTLDYDQYLETLEKVLTTVGEISKEDFTLYNHWSLSP----- 50
QY 60 MQYNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYOGGKGLKLLIDQL 119
DB 51 SIYHPYVITN-ASGIIVVATGMLFVEKKLIHECGKGVGHIEDISVAKSEGGKLGYYLVTSLS 109
QY 120 VTIGFDYGCYKIIILDCDEKNVFKYKCGFSNAGVEMQIR 158
DB 110 TKVAQENDCYKVIILDCSPENVGFYKCGYKGGVEMVCR 148
RESULT 6
ID ADU00602
XX AC ADU00602;
XX 13-JAN-2005 (first entry)
XX Amino acid sequence of a GNAL gene.
KW chitin; chitosan; fermentation;
KW glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;
KW glucosamine-6-phosphate synthase; GFAL gene; gms gene; chitin synthase;
KW glucosamine-6-phosphate acetyltransferase; GNAL gene; chitin synthase;

KW chitin deacetylase; CDA1 gene; CDA2 gene;
KW N-acetylglucosamine-6-phosphate deacetylase;
KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;
XX enzyme.
XX Candida albicans.
OS WO2004092391-A2.
PN 28-OCT-2004.
PD 12-APR-2004; 2004WO-US011286.
PF 11-APR-2003; 2003US-0462087P.
PR (ARKI-) ARKION LIFE SCI LLC.
PA Deng M, McMullin TW, Grund AD;
XX WPI; 2004-766880/75.
DR N-PSDB; ADU00601.
XX Producing chitin or chitosan, involves culturing microorganism with
PT genetic modification that results in increase in activity of glutamine-
PT fructose-6-phosphate amidotransferase, in fermentation medium, to produce
PT chitin or chitosan.
XX Disclosure; SEQ ID NO 35; 161pp; English.
XX The specification describes a method for producing chitin or chitosan by
CC a fermentation process. The method involves culturing in a fermentation
CC medium a microorganism which comprises one or more genetic modifications
CC that result in an increase in the activity of glutamine-fructose-6-
CC phosphate amidotransferase (also known as glucosamine synthetase and
CC glucosamine-6-phosphate synthase), and encoded by the GFAL eukaryotic gene
CC and the gms bacterial gene), glucosamine-6-phosphate acetyltransferase
CC (encoded by GNAL), chitin synthase or chitin deacetylase (encoded by CDA1
CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-
CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and
CC chitosanase, and collecting the chitin or chitosan. The method is useful
CC for producing chitin or chitosan by utilizing microorganisms such as
CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and
CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.
CC cerevisiae, A. niger or A. nidulans. The method enables high quantities
CC of chitin and chitosan to be produced cost effectively. The present
CC sequence represents a glucosamine-6-phosphate acetyltransferase, encoded
CC by a GNAL gene which may be used to transform microorganisms for use in the
CC method of the invention.
XX Sequence 149 AA;
Query Match 39.3%; Score 328; DB 8; Length 149;
Best Local Similarity 43.4%; Pred. No. 1.5e-27;
Matches 69; Conservative 26; Mismatches 52; Indels 12; Gaps 3;
QY 1 MSLPDGFYIRMEEGDLE-QVTEILKVLTTVTGTTTSPESFKLIKYWNEATVWNNEDKKI 59
DB 1 MMLPGQYTFRKLKLTLDYDQYLETLEKVLTTVGEISKEDFTLYNHWSLSP----- 50
QY 60 MQYNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYOGGKGLKLLIDQL 119
DB 51 SIYHPYVITN-ASGIIVVATGMLFVEKKLIHECGKGVGHIEDISVAKSEGGKLGYYLVTSLS 109
QY 120 VTIGFDYGCYKIIILDCDEKNVFKYKCGFSNAGVEMQIR 158
DB 110 TKVAQENDCYKVIILDCSPENVGFYKCGYKGGVEMVCR 148
RESULT 7
ID ABJ26242
XX ID ABJ26242 standard; protein; 190 AA.
XX AC ABJ26242;

XX	SQ	Sequence 174 AA;
		Query Match 25.9%; Score 216.5; DB 8; Length 174; Best Local Similarity 33.8%; Pred. No. 3.3e-15; Matches 53; Conservative 28; Mismatches 65; Indels 9; Gaps 3;
Qy	5	DGYIRMEGDLEQ-VTETLKVLTVTGTTSPESFKLYKWNATVWNNDEKKIMQYN 63 : : : : : : : : :
Db	27	DTIHRLECSHERGFALLSQLSPCLDITTSVFATRF-----AELAQDDHIILVAE 81 : : : : : : : :
Qy	64	PWIVDKRTEVTAATGNIIIERKIIIEHELGLCGHIEDIAVNASKYQGOGLGKLIDQLVTIG 123 : : : : : : : :
Db	82	DPASDER---ILATGCLFVERFLRGCKGVGHVEDVVVDAAARGGLGI RIVRLVEIS 138 : : : : : : : :
Qy	124	FDYGCYKIILDCEKNVKFYEKCFNSAGVEMOI 157 : : : : : : : :
Db	139	RDAACYKVIDCTPELRAYYAKGCFVEKGVMNAV 172 : : : : : : : :
RESULT 10		
AAU07779	ID	AAU07779 standard; protein; 184 AA.
XX	XX	AAU07779;
XX	AC	
XX	XX	(first entry)
DT	04-DEC-2001	
XX	Human novel transferase protein, NHP #22.	
DE	Human; transferase; breast cancer; prostate cancer; immunogen;	
XX	gene therapy; antisense.	
KW	Homo sapiens.	
XX		
OS		
XX		
FN	WO200164903-A2.	
XX		
PD	07-SEP-2001.	
XX		
PF	28-FEB-2001; 2001WC-US006460.	
XX		
PR	29-FEB-2000; 2000US-0185920P.	
XX		
PR	02-MAR-2000; 2000US-0186558P.	
XX		
PR	24-MAR-2000; 2000US-0191849P.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A;	
XX	Zambrowicz B, Sands AT, Walke DW, Wilganowski NL, Hu Y, Kieke JA;	
FI	Potter DG;	
XX		
DR	WPI; 2001-550185/61.	
XX	N-PSDB; AAS12627.	
DR		
PT	Noel nucleic acid sequences encoding novel human proteins useful for	
FT	diagnosis, drug screening, clinical trial monitoring and treatment of	
PT	diseases and disorders.	
XX		
PS	Claim 11; Page 59-60; 60pp; English.	
CC	The invention relates to isolated nucleic acids encoding novel human	
CC	transferase proteins [NHP]. The nucleic acids and proteins are useful for	
CC	diagnosis, drug screening, clinical trial monitoring and treatment of	
CC	diseases and disorders e.g. breast and prostate cancer. NHPs can also be	
CC	useful for augmenting the efficacy of chemotherapeutic agents used in	
CC	treatment of breast or prostate cancer. The nucleic acid is also useful	
CC	in NHP gene regulation, and as antisense primers in amplification	
CC	reactions of NHP gene sequences. NHPs are useful for producing	
CC	antibodies. The present sequence represents a novel human transferase	
XX		
SQ	Sequence 184 AA;	
	Query Match 25.0%; Score 209; DB 4; Length 184;	

Best Local Similarity 32.9%; Pred. No. 2 4e-14; Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

QY 5 DGFYIRMEGDLQEQ-VTETLKVLTGVTITPESFSLIKYWNATVWMDNEDKIMQYN 63
Db 37 EGLVLRLPLCTADLNRGFFKVLGQLTETGVVSPQFMKSPEHM-----KSGDYY 85

QY 64 PMWIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTTIG 123
Db 86 VTVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDCRGKQLKLLSTLTLLS 145

QY 124 FDYGCYKIILDCDEKNVKEKGFPS 149
Db 146 KKLNCYKITLECLPQNVGFYKKGVT 171

RESULT 11
AAB73505
ID AAB73505 standard; protein; 184 AA.
XX AAB73505;
XX
XX 31-JUL-2001 (first entry)
XX
XX Human transferase HTFS-12, SEQ ID NO:12.
XX
XX Human transferase; HTFS; agonist; antagonist; cellular signalling;
KW proliferation; cell proliferative disorder; immune disorder;
KW atherosclerosis; hepatitis; psoriasis; cancer; tumour; inflammation;
KW AIDS; Addison's disease; allergy; asthma; anaemia; cirrhosis;
KW Crohn's disease; atopic dermatitis; diabetes mellitus;
KW multiple sclerosis; rheumatoid arthritis; pancreatitis;
KW systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KW haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KW Gene therapy; drug screening.
XX
XX Homo sapiens.
XX
XX WO200132888-A2.
XX
XX 10-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-US030485.
XX
XX 04-NOV-1999; 99US-0163595P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih LL, Azimzai Y, Lu DAM, Baughn MR;
PI
XX WPI; 2001-328796/34.
DR N-PSDB; AAH23812.
XX
XX Human transferase polypeptides and polynucleotides useful for diagnosis,
PT prevention and treatment of cell proliferative and immune system
PT disorders and for identifying agonists and antagonists.
XX
XX Claim 1; Page 111; 157pp; English.
XX
XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HTFS proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferase activity, or for compounds that specifically bind to an HTFS
CC protein or which modulates the activity of an HTFS protein.
CC Pharmaceutical compositions comprising an HTFS protein. HTFS agonist or
CC antagonist, or genetic construct encoding an HTFS protein are useful for
CC treating a disease or condition associated with decreased or increased
CC expression of functional HTFS. Disorders which may be treated using such
CC compositions include cell proliferative disorders and immune disorders.
CC For example, diseases which may be treated include atherosclerosis,

CC hepatitis, psoriasis, cancers (including breast, bladder, bone marrow,
CC brain and uterus cancer), inflammation, AIDS, Addison's disease,
CC allergies, asthma, anaemia, cirrhosis, Crohn's disease, atopic
CC dermatitis, diabetes mellitus, multiple sclerosis, rheumatoid arthritis,
CC pancreatitis, systemic lupus erythematosus, thrombocytopenia, and
CC ulcerative colitis. They may also be used to treat complications of
CC cancer, haemodialysis, extracorporeal circulation, trauma and
CC haematopoietic cancer, including lymphoma, leukaemia and myeloma.
CC Polynucleotides encoding HTFS proteins are useful for creating transgenic
CC animals to model human diseases, for diagnostic purposes and to generate
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences. HTFS, and its catalytic or immunogenic fragments are useful
CC for screening libraries of compounds in a variety of drug screening
CC techniques. Antibodies which specifically bind HTFS may be used for the
CC diagnosis of disorders associated with the expression of HTFS, or in
CC assays to monitor patients being treated with HTFS or agonists,
CC antagonists or inhibitors of HTFS. The present sequence represents an
CC HTFS protein of the invention
XX
XX Sequence 184 AA;
SQ

Query Match 24.9%; Score 208; DB 4; Length 184;
Best Local Similarity 32.9%; Pred. No. 3.1e-14;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

QY 5 DGFYIRMEGDLQEQ-VTETLKVLTGVTITPESFSLIKYWNATVWMDNEDKIMQYN 63
Db 37 EGLVLRLPLCTADLNRGFFKVLGQLTETGVVSPQFMKSPEHM-----KSGDYY 85

QY 64 PMWIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTTIG 123
Db 86 VTVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDCRGKQLKLLSTLTLLS 145

QY 124 FDYGCYKIILDCDEKNVKEKGFPS 149
Db 146 KKLNCYKITLECLPQNVGFYKKGVT 171

RESULT 12
AAG67123
ID AAG67123 standard; protein; 184 AA.
XX
XX AAG67123;
XX
XX 13-NOV-2001 (first entry)
XX
XX Amino acid sequence of human 50287 transferase.
XX
XX Human; transferase; 25324 transferase; 50287 transferase; cancer;
KW 28899 transferase; 47007 transferase; 42967 transferase; angiogenesis;
KW kynurenine aminotransferase; phosphatidyl transferase; brain tumour;
KW cell proliferation; cell differentiation.
XX
XX Homo sapiens.
XX
XX WO200164904-A2.
XX
XX 07-SEP-2001.
XX
XX 27-FEB-2001; 2001WO-US006463.
XX
XX 29-FEB-2000; 2000US-0185711P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R, Macbeth KJ, Rudolph-Owen L;
PI
XX WPI; 2001-550186/51.
DR N-PSDB; AAH75151.
XX
XX Novel transferase polypeptides and polynucleotides encoding the
PT polypeptides, useful for the diagnosis or the identification of a
PT compound capable of treating cancer or a disorder characterized by

```

PT aberrant angiogenesis.
XX
PS Claim 4; Fig 4; 185pp; English.
XX
CC The present sequence represents a human transferase polypeptide. The
CC specification describes 25324, 50287, 28899, 47007 and 42967
CC transferases. The 25324 transferase is homologous to kynurenine
CC aminotransferase. The 47007 transferase is homologous to phosphatidyl
CC transferase. The transferase polypeptides and polynucleotides may be used
CC for the diagnosis or the identification of a compound capable of treating
CC cancer (e.g., lung, breast, or colon) or a disorder characterised by
CC aberrant angiogenesis (e.g., brain tumour angiogenesis). Modulators of
CC the transferases may be used to treat the above, or modulate cellular
CC proliferation and/or differentiation, or a subject at risk of the above
CC disorders
XX
SQ Sequence 184 AA;
    Query Match          24.9%; Score 208; DB 4; Length 184;
    Best Local Similarity 32.9%; Pred. No. 3.1e-14;
    Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;
QY 5 DGFYIRMEGDLQ-VTETLKVLTVTGTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63
DB 37 EGLVLRPLCTADLNRFKVLGQLTETGVVSPEQFMKSFEHM-----KKS GDY 85
QY 64 PMVIVDKRTETVAATGNIIEERKIIHELGLCHIEDAVNSKYOGGLKLLIDQLVTIG 123
DB 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRGRVEDVVVSDECRGKQLGLLSTLTLLS 145
QY 124 FDYGCYKIILDCDEKNVKFYKCGFS 149
DB 146 KKLNCYKITLECLPQNVGFKFGYT 171

RESULT 13
ADRI0847
XX ADRI0847 standard; protein; 184 AA.
XX
AC ADRI0847;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human FLJ10607, modifier of axin pathway.
XX
KW Human; FLJ10607; axin; cytosstatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 89..171
FT /note= "Acetyltransferase (GNAT) domain"
XX
PN WO2004061086-A2.
XX
PD 22-JUL-2004.
XX
PF 29-DEC-2003; 2003WO-US041619.
XX
PR 30-DEC-2002; 2002US-0436965P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K;
XX
WPI; 2004-534377/51.
DR N-PSDB; ADRI0844, ADRI0846.
DR REFSEQ; XM_085119.
XX
PT Use of FLJ10607 polypeptide or nucleic acid for identifying a candidate
PT Axin pathway modulating agent, modulating an Axin pathway of a cell, or
PT diagnosing a disease, e.g. cancer in a patient.
XX

```

```

PS Disclosure; SEQ ID NO 4; 55pp; English.
XX
CC The present sequence is that of a human FLJ10607 polypeptide. This is the
CC human orthologue of a Caenorhabditis elegans protein that has been shown
CC to have the ability to modify the axin pathway. The invention provides
CC methods for using FLJ10607 axin modifier genes and polypeptides to
CC identify FLJ10607-modulating agents that are candidate therapeutic agents
CC useful for the diagnosis and treatment of disorders associated with
CC defective or impaired axin function and/or FLJ10607 function, especially
CC proliferative disorders such as cancer. Preferred FLJ10607-modulating
CC agents specifically bind to FLJ10607 polypeptides and restore axin
CC function, or are nucleic acid modulators such as antisense oligomers and
CC RNA interference (RNAi) agents that repress FLJ10607 gene expression or
CC product activity, e.g. by binding to and inhibiting the respective
CC nucleic acid (i.e. DNA or mRNA). Candidate FLJ10607 modulating agents are
CC tested with an assay system comprising a FLJ10607 polypeptide or nucleic
CC acid.
XX
SQ Sequence 184 AA;
    Query Match          24.9%; Score 208; DB 8; Length 184;
    Best Local Similarity 32.9%; Pred. No. 3.1e-14;
    Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;
QY 5 DGFYIRMEGDLQ-VTETLKVLTVTGTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63
DB 37 EGLVLRPLCTADLNRFKVLGQLTETGVVSPEQFMKSFEHM-----KKS GDY 85
QY 64 PMVIVDKRTETVAATGNIIEERKIIHELGLCHIEDAVNSKYOGGLKLLIDQLVTIG 123
DB 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRGRVEDVVVSDECRGKQLGLLSTLTLLS 145
QY 124 FDYGCYKIILDCDEKNVKFYKCGFS 149
DB 146 KKLNCYKITLECLPQNVGFKFGYT 171

RESULT 14
ADRI0848
XX ADRI0848 standard; protein; 184 AA.
XX
AC ADRI0848;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human FLJ10607, modifier of axin pathway.
XX
KW Human; FLJ10607; axin; cytosstatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 89..171
FT /note= "Acetyltransferase (GNAT) domain"
XX
PN WO2004061086-A2.
XX
PD 22-JUL-2004.
XX
PF 29-DEC-2003; 2003WO-US041619.
XX
PR 30-DEC-2002; 2002US-0436965P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K;
XX
WPI; 2004-534377/51.
DR N-PSDB; ADRI0845.
DR REFSEQ; NM_198066.2.
XX
PT Use of FLJ10607 polypeptide or nucleic acid for identifying a candidate
PT Axin pathway modulating agent, modulating an Axin pathway of a cell, or

```

PT diagnosing a disease, e.g. cancer in a patient.
 XX
 PS Example 1; SEQ ID NO 5; 55pp; English.
 XX
 CC The present sequence is that of a human FLJ10607 polypeptide. This is the
 CC human orthologue of a Caenorhabditis elegans protein that has been shown
 CC to have the ability to modify the axin pathway. The invention provides
 CC methods for using FLJ10607 axin modifier genes and polypeptides to
 CC identify FLJ10607-modulating agents that are candidate therapeutic agents
 CC useful for the diagnosis and treatment of disorders associated with
 CC defective or impaired axin function and/or FLJ10607 function, especially
 CC proliferative disorders such as cancer. Preferred FLJ10607-modulating
 CC agents specifically bind to FLJ10607 polypeptides and restore axin
 CC function, or are nucleic acid modulators such as antisense oligomers and
 CC RNA interference (RNAi) agents that repress FLJ10607 gene expression or
 CC product activity, e.g. by binding to and inhibiting the respective
 CC nucleic acid (i.e. DNA or mRNA). Candidate FLJ10607 modulating agents are
 CC tested with an assay system comprising a FLJ10607 polypeptide or nucleic
 CC acid.
 XX
 SQ Sequence 184 AA;

Query Match 24.9%; Score 208; DB 8; Length 184;
 Best Local Similarity 32.9%; Pred. No. 3.1e-14;
 Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;
 QY 5 DGFYIRMEEGDLEQ-VTETLKVLTGVTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63
 Db 37 EGLVLRPLCTADLNRGFFKVLGQLTETGVVSPEQFMKSFHM-----KKS GDY 85
 QY 64 PMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTIG 123
 Db 86 VTVVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDECRGKQLKLLSTLTLLS 145
 QY 124 PDYGYCKIILDCDEKNVFKYERKGF 149
 Db 146 KKLNCYKITLECLPQNVGFYKFGYT 171

RESULT 15
 AAB63700
 ID AAB63700 standard; protein; 200 AA.
 AC AAB63700;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human gastric cancer associated antigen protein sequence SEQ ID NO:1062.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW Cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 WP 2001-025274/03.
 XX
 PT Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.

XX
 PS Example 1; Page 675; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytosolic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 SQ Sequence 200 AA;
 Query Match 24.9%; Score 208; DB 4; Length 200;
 Best Local Similarity 32.9%; Pred. No. 3.4e-14;
 Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;
 QY 5 DGFYIRMEEGDLEQ-VTETLKVLTGVTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63
 Db 53 EGLVLRPLCTADLNRGFFKVLGQLTETGVVSPEQFMKSFHM-----KKS GDY 101
 QY 64 PMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTIG 123
 Db 102 VTVVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDECRGKQLKLLSTLTLLS 161
 QY 124 PDYGYCKIILDCDEKNVFKYERKGF 149
 Db 162 KKLNCYKITLECLPQNVGFYKFGYT 187
 Search completed: June 14, 2006, 15:21:07
 Job time : 78.2383 secs

THIS PAGE IS BLANK

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE IS BLANK

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE IS BLANK